

# Learning Factor Models of Students at Risk of Failing in the Early Stage of Tertiary Education

Geraldine Gray  
Colm McGuinness  
Philip Owende  
Markus Hofmann

Institute of Technology Blanchardstown, Ireland  
[geraldine.gray@itb.ie](mailto:geraldine.gray@itb.ie)

**ABSTRACT:** This paper reports on a study to predict students at risk of failing based on data available prior to commencement of first year. The study was conducted over three years, 2010 to 2012, on a student population from a range of academic disciplines,  $n=1,207$ . Data was gathered from both student enrollment data and an online, self-reporting, learner-profiling tool administered during first-year student induction. Factors considered included prior academic performance, personality, motivation, self-regulation, learning approaches, age, and gender. Models were trained on data from the 2010 and 2011 student cohort, and tested on data from the 2012 student cohort. A comparison of eight classification algorithms found  $k$ -NN achieved best model accuracy (72%), but results from other models were similar, including ensembles (71%), support vector machine (70%), and a decision tree (70%). However, improvements in model accuracy attributable to non-cognitive factors were not significant. Models of subgroups by age and discipline achieved higher accuracies, but were affected by sample size;  $n < 900$  underrepresented patterns in the dataset. Factors most predictive of academic performance in first year of study at tertiary education included age, prior academic performance, and self-efficacy. Early modelling of first-year students yielded informative, generalizable models that identified students at risk of failing.

**Keywords:** Learning analytics, learner profiling, academic performance, non-cognitive factors of learning, tertiary education

## 1 INTRODUCTION

Enrollment numbers to tertiary education are increasing, as is diversity in student populations (OECD, 2013; Patterson, Carroll, & Harvey, 2014); however, significant numbers of students do not complete the courses in which they enroll, particularly courses with lower entry requirements (ACT, 2012; Mooney, Patterson, O'Connor, & Chantler, 2010). Factors predictive of academic performance have been the focus of research for many years (Farsides & Woodfield, 2003; Moran & Crowley, 1979), and continue as an active research topic (Jayaprakash, Moody, Lauria, Regan, & Baron 2014; Cassidy, 2011; Wise & Shaffer, 2015), indicating the inherent difficulty in generating accurate learning factor models (Knight, Buckingham Shum, & Littleton, 2013; Tempelaar, Cuypers, van de Vrie, Heck, & van der Kooij, 2013).

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

Tertiary education providers collect much data on students, including demographic data, academic activity, and log data from online campus activities. As a result, the application of data analytics to educational settings is an emerging and growing research discipline of data analytics (Campbell, deBlois, & Oblinger, 2007; Mirriahi, Gašević, Long, & Dawson, 2014; Sachin & Vijay, 2012; Siemens & Baker, 2012). The primary aim of learning analytics is to provide learning professionals, and students, with actionable information that can be used to enhance the learning process (Siemens, 2012; Chatti, Dyckhoff, Schroeder, & Thüs, 2012). Much of the published work in learning analytics is based on data systematically gathered by education providers, particularly log data from Virtual Learning Environments (VLE) and Intelligent Tutoring Systems (ITS) (Drachsler & Greller, 2012; Baker & Yacef, 2010). Measurement and analysis of non-cognitive factors of learning has received limited attention from the learning analytics community (Buckingham Shum & Deakin Crick, 2012) with the exception of factors inferred from online behaviour (e.g., Ali, Hatala, Gašević, & Winne, 2014; Arnold & Pistilli, 2012; Shute & Ventura, 2013). In contrast, many publications emanating from educational psychology report on correlations with, or dependencies on, a wide variety of non-cognitive psychometric factors of learning and academic performance (Dekker, Pechenizkiy, & Vleeshouwers, 2009; Herzog, 2006; Robbins, Lauver, Le, Davis, & Langley, 2004). Therefore, further work is needed to determine the value of non-cognitive factors in models of learning over and above the analysis of data systematically gathered by education providers.

Early identification and profiling of students at risk of failing informs provisioning of supports and modifications to learning environments to enable all students perform optimally. Colby (2005) identified week two as a critical point in identifying at-risk students. Arnold and Pistilli (2012) reported positive results identifying at-risk students in week two based on interactions with a VLE, prior academic performance, and demographic data. Similarly, Milne, Jeffrey, Suddaby, and Higgins (2012) reported successful results in predicting students at risk of failing based on analysis of online behaviour in week one, while early results reported in Gray, McGuinness, and Owende (2014) suggested non-cognitive factors of learning measured during student induction could also achieve good predictive accuracy.

The objective of this study was to investigate if factors of learning, measured prior to commencement of first year, could accurately predict students at risk of failing at the end of year one of tertiary education. The data used covered a diverse student population enrolled in a range of academic courses.

## 2 FACTORS OF LEARNING

Research from educational psychology has identified a range of factors related directly or indirectly to academic performance in tertiary education. In particular, factors of prior academic performance, and non-cognitive factors of personality, motivation, self-regulation, and approaches to learning are cited as predictive of first-year GPA (Cassidy, 2011; Chamorro-Premuzic & Furnham, 2008; Gray, McGuinness, Owende, & Carthy, 2014; Komarraju & Nadler, 2013; Robbins et al., 2004; Swanberg & Martinsen, 2010). In addition, a number of non-cognitive factors of learning have been associated with an effective learning disposition, including a deep learning approach, ability to self-regulate, setting learning goals,

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

persistence, conscientiousness, and sub-factors of openness, namely intellectual curiosity, creativity, and open-mindedness (Buckingham Shum & Deakin Crick, 2012; Knight et al., 2013; Tishman, Jay, & Perkins, 1993).

Prior academic performance can refer to high school grade point average (HSGPA) or equivalent, and/or standardized tests such as American College Testing (ACT) scores or Scholastic Aptitude Test (SAT) scores. Correlations ( $r$ ) between prior academic performance and academic performance in tertiary education are consistent and relatively strong across studies of younger college students. For example, a meta-analysis of 109 studies conducted by Robbins et al. (2004) found a moderate correlation between academic performance and HSGPA ( $r=0.448$ , 90% CI [0.41, 0.49]). However, Eppler and Harju (1997) found that correlations between academic performance and SAT scores were not as strong for mature students<sup>1</sup> ( $r=0.09$ , 90% CI\* [-0.15, 0.32],  $n=50$ ).

There is broad agreement that conscientiousness is the best personality-based predictor of academic performance (Swanberg & Martinsen, 2010). For example, Komarraju, Karau, Schmeck, and Avdic (2011) reported a correlation of  $r=0.29$  ( $p<0.01$ , 90% CI\* [0.2, 0.37],  $n=308$ ) between conscientiousness and academic performance. Some studies have also cited openness as significant; however, reported correlations with academic performance are less consistent (Gray, McGuinness, Owende, & Carthy, 2014). For example, Komarraju et al. (2011) reported a correlation of  $r=0.13$  ( $p<0.05$ , 90% CI\* [0.04, 0.22],  $n=308$ ) between openness and GPA, while Kappe and van der Flier (2010) reported a correlation of  $r=-0.08$  (90% CI\* [-0.22, 0.06],  $n=133$ ). Differences may be explained by assessment methods, as open personalities tend to do better when assessment methods are unconstrained by submission rules and deadlines (Kappe & van der Flier, 2010).

Motivation to learn is characterized by a range of complementary theories (Steel & Konig, 2006). A meta-analysis of 109 studies analyzing psychosocial and study skill traits found two factors of motivation, namely self-efficacy ( $r=0.495$ , 90% CI [0.44, 0.55]) and achievement motivation ( $r=0.303$ , 90% CI [0.26, 0.34]), had the highest correlations with academic performance (Robbins et al., 2004). Achievement motivation may reflect a goal to succeed or a desire to avoid failure (Elliot & Harackiewicz, 1996). Achievement goals in turn can be assessed with respect to learning (intrinsic) achievement and/or performance (extrinsic) achievement (Eppler & Harju, 1997). Studies have found that learning goals have a higher correlation with academic performance than performance goals (Gray, McGuinness, Owende, & Carthy, 2014). For example, Eppler and Harju (1997) reported a correlation of  $r=0.3$  ( $p<0.001$ , 90% CI\* [0.19, 0.4],  $n=212$ ) for learning goals and  $r=0.13$  ( $p=0.059$ , 90% CI\* [0.02, 0.24],  $n=212$ ) for performance goals. Covington (2000) argued goal setting in itself was not enough, as ability to self-regulate learning can be the difference between achieving, or not achieving, the goals set. However, Komarraju and Nadler (2013) found that self-regulation (monitoring and evaluating learning)

---

<sup>1</sup> CI\* denotes confidence intervals were not provided by the author, and were calculated in R version 3.0.2 using CIs in package *psychometric*, which calculates confidence intervals based on the Fisher  $r$ -to- $z$  transformation.

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

did not account for any additional variance in academic performance over and above self-efficacy, but study effort and study time did account for additional variance.

A number of studies reported the relationship between academic performance and temperament or motivation was mediated by approach to learning task (Bruinsma, 2004; Chamorro-Premuzic & Furnham, 2008; Diseth, 2011; Sins, van Joolingen, Savelsbergh, & van Hout-Wolters, 2008). Marton and Säljö (2005) classified a student's approach to learning as either shallow or deep. Deep learners aim to understand content, while shallow learners aim to memorize content regardless of their level of understanding. Later studies added strategic learners (Entwhistle, 2005, p. 19), whose priority is to do well, and will adopt either a shallow or a deep learning approach depending on the requisites for academic success. Comparing the influence of approaches to learning on academic performance, Cassidy (2011) found correlations with GPA were similar for both the strategic ( $r=0.32$ ,  $p<0.01$ , 90% CI\* [0.16, 0.46],  $n=97$ ) and the deep ( $r=0.31$ ,  $p<0.01$ , 90% CI\* [0.15, 0.45],  $n=97$ ) learning approach. However, Chamorro-Premuzic and Furnham (2008) found a deep learning approach ( $r=0.31$ ,  $p<0.01$ , 90% CI [0.19, 0.42],  $n=158$ ) had a higher correlation with GPA than a strategic learning approach ( $r=0.18$ ,  $p<0.05$ , 90% CI\* [0.05, 0.3],  $n=158$ ). Differences found have been explained, in part, by assessment type (Volet, 1996), highlighting the importance of assessment design in encouraging effective learning strategies.

As evidenced by results cited in this section, statistical models have dominated data analysis in educational psychology (Dekker et al., 2009), particularly Pearson correlation and linear regression. There is evidence to suggest that assumptions for Pearson correlation and linear regression are not always met when analyzing non-cognitive factors of learning (Gray, McGuinness, Owende, & Carthy, 2014). Therefore, it is pertinent to ask if data mining's empirical modelling approach is more appropriate for models of learning. A growing number of learning analytics studies has investigated the role of non-cognitive factors in models of learning (Arnold & Pistilli, 2012; Buckingham Shum & Deakin Crick, 2012; Shute & Ventura, 2013). For example, Bergin (2006) cited an accuracy of 82% using an ensemble model predicting a strong ( $\text{grade}>55\%$ ) or weak ( $\text{grade}<55\%$ ) performance based on prior academic performance, self-efficacy, and study hours ( $n=58$ ). Gray, McGuinness, and Owende (2013) cited similar accuracies (81%,  $n=350$ ) with a Support Vector Machine model using cognitive and non-cognitive factors of learning to distinguish high-risk ( $\text{GPA}<2.0$ ) from low-risk ( $\text{GPA}\geq 2.5$ ) students based on first-year GPA.

### 3 THE STUDY DATASET

#### 3.1 Description of Study Participants

The study participants were first-year students at the Institute of Technology Blanchardstown (ITB), Ireland. The admission policy at ITB supports the integration of a diverse student population in terms of age and socio-economic background. Course entry requirements are generally lower than corresponding university courses (Mooney et al., 2010).

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

Each September from 2010 to 2012, all full-time, first-year students at ITB were invited to participate in the study by completing learner profiling administered during first-year student induction. A total of 1,365 (52%) full-time, first-year students completed the profiling. Eliminating invalid student ID's (n=100), those who did not give permission to be included in the study (n=35), students under the age of 18 at the time of profiling (n=3), and instances missing profiling data (n=20) resulted in 46% of full-time, first-year students participating in the study (n=1,207).

Participants ranged in age from 18 to 60, with a mean age of 23.27 (standard deviation,  $s=7.3$ ); of which 355 (29%) students were mature (23 and over<sup>2</sup>), 713 (59%) were male, and 494 (41%) were female. Students were enrolled in a range of courses in the disciplines of Business (n=402, 33%), Humanities (n=353, 29%), Information Technology (n=239, 20%), Engineering (n=172, 14%), and Horticulture (n=41, 3%).

### 3.2 Study Factors and Instruments Used

The study dataset included data from three sources: 1) student registration; 2) non-cognitive factors of learning measured during first-year induction; and 3) exam results from the first year of study at ITB, supplied by the college. Study factors are in italics in following sections.

#### 3.2.1 Student registration data

Registration data included *age*, *gender*, and prior academic performance. Access to full-time college courses in Ireland is based on academic performance in the leaving certificate (or equivalent), a set of state exams at the end of secondary school. The leaving certificate includes four mandatory subjects, namely mathematics, English,<sup>3</sup> Irish, and a foreign language, and typically an additional three elective subjects. Subjects can be studied at higher or ordinary level; mathematics and Irish are also offered at the foundation level. College places are offered based on CAO<sup>4</sup> points, an aggregate score based on points achieved in a student's best six leaving certificate subjects (range [0,600]). Table 12 in the Appendix maps exam grades to CAO points for higher level, ordinary level, and foundation level examinations. CAO points are awarded for marks  $\geq 40\%$ .

The study dataset included *CAO points*, points in *mathematics*, and points in *English* for each student. Points achieved in additional subjects (53 in total) were included as average points achieved by subject category. The Department of Education in Ireland groups leaving certificate subjects into six categories

---

<sup>2</sup> This is a state-wide definition of a mature student; their entry requirements are less strict.

<sup>3</sup> The Leaving Certificate English syllabus aims to develop a mature and critical literacy, a respect and appreciation for language, and an awareness of the value of literature ([www.education.ie](http://www.education.ie)).

<sup>4</sup> CAO refers to the Central Applications Office with responsibility for processing applications for undergraduate courses in the Higher Education Institutes in Ireland.

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

based on subject content, namely humanities, social, artistic, practical, science, and business.<sup>5</sup> These were combined to create three categories as follows: *applied* (artistic and practical categories); *humanities* (humanities and social categories); and *methodical* (science and business categories). Table 1 lists subjects included in each category. Many science and business subjects are numerate based; however, the two most popular electives in the methodical category did not have a significant mathematics component, namely Biology (51% of students) and Business (48% of students). Electives from the humanities category were the most popular; for example, 56% studied Geography and 51% studied French. All subjects in the applied category had a significant practical component; however, 43% of participants did not have a grade for this category of electives, limiting its usefulness.

Descriptive statistics for study factors of prior academic performance in Table 2 confirmed a student population with a weaker prior academic profile compared to university students as reported in Mooney et al. (2010). Of particular note was the low average points in mathematics (mean,  $m=23.8$ , equivalent to 55%–59% in an ordinary level paper), which was significantly lower than all other subject areas.

**Table 1: Leaving certificate subject categories.**

Category	Description
Applied	Art, Building Construction, Craft Design and Technology, Engineering, Graphic and Tech Design, Music, Music and Musicianship, Drama and Theatre Studies, Technical Drawing, Technology, Leaving Certificate Link Modules
Humanities	Classical Studies, Economic and Social History, English, Geography, History, Home Economics, all Languages, Religious Studies
Methodical	Accounting, Agricultural Economics, Agricultural Science, Applied Mathematics, Biology, Business (Organization/Studies), Chemistry, Computer (Science/Studies), Economics, Mathematics, Physics, Physics with Chemistry, Science/Environment Science

**Table 2: Leaving certificate points by subject category ( $m \pm s$ ).**

Subject	Average Points	Subject Category	Average Points
CAO points	259.5 $\pm$ 78.1 (n=1,018)	Applied Average	48.5 $\pm$ 19.5 (n=647)
English	46.4 $\pm$ 18.5 (n=1,015)	Humanities Average	40.0 $\pm$ 14 (n=1,016)
Mathematics	23.8 $\pm$ 13.9 (n=1,008)	Methodical Average	32.1 $\pm$ 15.5 (n=1,016)

Valid range for CAO points is [0,600]. Range for subjects and subject categories is [0,100]. Means ( $m$ ) and standard deviations ( $s$ ) were calculated based on the number of participants who had results in each category, as indicated by “n” above.

### 3.2.2 Additional non-cognitive factors gathered

The following sections discuss fifteen non-cognitive factors of learning included in the study, and the

<sup>5</sup> Details of subject groups can be found at the Department of Education’s National Career Guidance website: [www.careersportal.ie](http://www.careersportal.ie)



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

profiling tool used to gather the data. Questionnaire items were primarily taken from validated instruments in the public domain and administered during first-year student induction using an online tool developed for this study ([www.howilearn.ie](http://www.howilearn.ie)). The wording of some questions was changed to suit the context; for example, *Follow a schedule* was changed to *I like to do things according to a plan or schedule*. Unless otherwise stated, items used a five-level Likert scale. Since questionnaire length can affect the quality of response (Burisch, 1997; Galesic & Bosnjak, 2009), the number of items was reduced for some scales by removing similar items despite the likely negative impact on internal reliability statistics. Table 3 gives descriptive statistics for each factor.

The personality factors included were *conscientiousness* and *openness*. Items for both scales were taken from the International Personality Item Pool (IPIP) scales (Goldberg et al., 2006). Six items were selected from the Conscientiousness Big-Five Domain scale, and six items were selected from the Openness to Experience, NEO Domain scale.

Motivation was assessed based on *self-efficacy*, and two achievement motivation scales, *intrinsic* (learning) *goal orientation* and *extrinsic* (performance) *goal orientation*. Scales were from the Motivation Strategies for Learning Questionnaire (MSLQ) (Pintrich, Smith, Garcia, & McKeachie, 1991). All four items from each of the intrinsic goal-orientation and extrinsic goal-orientation scales were included. Three of the eight items from the self-efficacy scale were included.

Three factors of self-regulation were included: 1) *metacognitive self-regulation*, 2) study time and environment (*study time*), and 3) effort regulation (*study effort*). Scales were from MSLQ. To facilitate administration during student induction, items were selected based on their relevance to prior academic experiences. Five items were included from the twelve-item metacognitive self-regulation scale, four items were included from the eight-item time and study environment scale, and three items were included from the four-item effort regulation scale.

**Table 3: Non-cognitive factors of learning.**

Category & Instrument	Factor	m ± s	95% CI
Personality, Goldberg's IPIP scales ( <a href="http://ipip.ori.org">http://ipip.ori.org</a> )	Conscientiousness	5.95 ± 1.53	[5.86, 6.03]
	Openness	6.07 ± 1.29	[5.99, 6.14]
Motivation, based on MSLQ (Pintrich et al., 1991)	Self-efficacy	6.85 ± 1.42	[6.77, 6.93]
	Intrinsic goal orientation	7.09 ± 1.36	[7.03, 7.17]
	Extrinsic goal orientation	7.81 ± 1.38	[7.73, 7.89]
Self-regulated learning, based on MSLQ (Pintrich et al., 1991)	Metacognitive self-regulation	5.88 ± 1.36	[5.80, 5.95]
	Study effort	5.93 ± 1.77	[5.83, 6.03]
	Study time	6.17 ± 2.32	[6.04, 6.30]
Learning style, based on R-SPQ-2F (Biggs, Kember, & Leung, 2001)	Deep learner	5.36 ± 2.91	[5.20, 5.53]
	Shallow learner	1.33 ± 1.95	[1.22, 1.44]
	Strategic learner	3.41 ± 2.48	[3.27, 3.55]

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372.  
<http://dx.doi.org/10.18608/jla.2016.32.20>

Category & Instrument	Factor	m ± s	95% CI
Preferred learning channel, NLN Learning Styles Questionnaire.	Visual	7.17 ± 2.06	[7.05, 7.28]
	Auditory	3.13 ± 2.17	[3.04, 3.29]
	Kinaesthetic	4.67 ± 2.42	[4.53, 4.80]
	Preference for group work	6.55 ± 3.36	[6.36, 6.74]

m=mean; s=standard deviation. Valid range for each factor is [0,10].

Learning approach was assessed based on the Revised Two-Factor Study Process Questionnaire (R-SPQ-2F) published by Biggs et al. (2001). The published questionnaire provided separate scales for *shallow* and *deep learning approaches*. Items used a five-level Likert scale. The question style was changed for this study, forcing participants to choose between a *deep, strategic, or shallow learning approach*. Each item on a four-item scale asked participants to pick one of three statements: two statements, relating to *deep* and *shallow learning approach*, were taken from R-SPQ-2F; the third statement, relating to a *strategic learning approach*, was compiled in collaboration with the National Learning Network Assessment Service<sup>6</sup> (NLN). The style of question matched the style of items on a learning-styles profiler designed by NLN and used by ITB in previous years.

In agreement with NLN, scales from their learning styles questionnaire were also included. This covered *learner modality* (*Visual, Auditory* and/or *Kinaesthetic* [VAK]; Fleming, 1995), which was scored from six questions, each offering two choices of modality, resulting in four items per modality across the six questions. Preference for solo or *group work* was also asked.

Questionnaire validity and internal reliability were assessed using a paper-based questionnaire that included both the revised wording of questions used on the online questionnaire (reduced scale), and the original questions from the published instruments (original scale). The paper questionnaire was administered during scheduled first-year lectures across all academic disciplines. Results are detailed in Table 4. Pearson correlations between scores calculated from the reduced scale, and scores calculated from the original scale, were high for all factors ( $\geq 0.9$ ) except *intrinsic goal orientation* and *study time*. Internal reliability was assessed using Cronbach's Alpha. All factors had acceptable reliability ( $>0.7$ ) given the small number of questions per scale<sup>7</sup> (between 3 and 6), with the exception again of *intrinsic goal orientation* and *study time*. Interestingly, Komarraju and Nadler (2013) reported similar difficulties with the intrinsic goal orientation scale when administered in the first week of term. *Intrinsic goal orientation* and *study time* were not removed from the dataset; however, it is acknowledged that inferences based on these factors may be unreliable.

<sup>6</sup> The National Learning Network Assessment Service provides functional strategies and support for children, adolescents, and adults with specific learning difficulties. They are located on campus at ITB ([www.nln.ie](http://www.nln.ie)).

<sup>7</sup> While generally a Cronbach Alpha of  $>0.8$  indicates good internal consistency, Cronbach Alpha closer to 0.7 can be regarded as acceptable for scales with fewer items (Tavakol & Dennick, 2011).



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372.  
<http://dx.doi.org/10.18608/jla.2016.32.20>

**Table 4: Questionnaire validation: correlations and Cronbach's Alpha.**

Factor	n	Reduced scale		Original scale		Correlations between original scale and reduced scale scores	Published $\alpha$
		$\alpha$	items	$\alpha$	items		
Openness	47	0.70	6	0.84	10	0.90 (95% CI [0.82, 0.94])	0.82
Conscientiousness	42	0.69	6	0.80	10	0.95 (95% CI [0.91, 0.97])	0.79
Intrinsic Goals	43	0.63	4	0.53	4	0.81 (95% CI [0.68, 0.89])	0.74
Extrinsic Goals	48	0.69	4	0.58	4	0.90 (95% CI [0.82, 0.94])	0.62
Effort	41	0.69	3	0.74	4	0.98 (95% CI [0.96, 0.99])	0.69
Self-efficacy	48	0.82	3	0.81	7	0.93 (95% CI [0.89, 0.97])	0.94
Meta Cognitive							
Self-regulation	38	0.70	5	0.70	12	0.90 (95% CI [0.81, 0.95])	0.79
Time and study	48	0.55	4	0.68	8	0.79 (95% CI [0.65, 0.87])	0.76
Learning Style	42	0.76	4				

### 3.2.3 Year 1 academic performance

First-year academic performance was measured as Grade Point Average (GPA), an aggregate score of between 10 and 12 first-year modules (range [0,4]). GPA is calculated from module grade achieved multiplied by module weighting (credits). A GPA<2.0, or a result of *fail* in any individual module, results in an award of *Fail* overall, otherwise a student is awarded a *Pass* result and may progress to the next academic stage. Table 5 shows the academic profile of study participants across GPA bands. Of the students with GPA $\geq$ 2.5 (n=558, 46%), 92% passed all modules indicating a low-risk group that can progress to year two. Of the students with GPA<2 (n=432, 36%), 91% failed three or more modules, indicating a high-risk group falling well short of progression requirements. For the remaining students in the GPA band<sup>8</sup> [2.0, 2.5) (n=217, 18%), 35% passed all modules, 36% failed one module, 20% failed two modules, and 8% failed more than two modules. This is a less homogenous group in terms of academic results, but could be regarded as struggling academically (medium-risk), either passing all modules with low grades or required to repeat one or two modules to progress.

**Table 5: Number of modules passed, by GPA band.**

GPA band	n	Failed > 6 modules	Failed 3 to 6 modules	Failed 1 to 2 modules	Passed all modules
0.0≤GPA<2.0	432	245 (57%)	146 (34%)	39 (9%)	1 (0.2%)
2.0≤GPA<2.5	217	0 (0%)	18 (8%)	122 (56%)	77 (35%)
2.5≤GPA≤4	558	0 (0%)	5 (1%)	38 (7%)	515 (92%)

## 4 ANALYTICAL APPROACHES USED

Both statistical analysis and classification techniques were used in this study. Correlation analysis facilitated comparison with other studies in educational psychology. Analysis of subgroup differences by

<sup>8</sup> [x,y) denotes a range inclusive of x but exclusive of y.

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

GPA band, age group, and gender aided interpretation of classification model results.

#### 4.1 Correlation Analysis

Pearson product-moment correlation coefficients ( $r$ ) were calculated between all study factors and GPA. An assumption of calculating the significance of a Pearson's correlation is that attributes are normally distributed. However, all study attributes failed a Shapiro-Wilk normality test ( $p > 0.05$ ). This is common in data relating to education and psychology (Kang & Harring, 2015; Micceri, 1989), which is likely to be skewed, have a heavy or light tail, and/or be multimodal (Smith & Wells, 2006). Therefore, significance was verified using 1,999 bootstrap confidence intervals (B-CI) using the bias corrected and accelerated method (Carpenter & Bithell, 2000) as implemented in R version 3.0.2.

#### 4.2 Analysis of Group Differences

Group differences were assessed for subgroups by GPA band (three groups), age category (three groups), and gender (two groups). Student's t-test assumes a normal distribution and equality of variance for the test-statistic attribute mean. Applying a Shapiro-Wilk test to attribute means of 50 bootstrap samples confirmed normality. However, a Brown-Forsythe test to compare variance in attribute means found variances were not equal. Therefore, results from Welch's t-test are reported for group differences by gender. One-way Analysis of Variance (ANOVA) with post hoc Tukey tests was used for multiple groups. ANOVA is robust regarding assumptions of equality of variance except in extreme cases (Hair, Black, Babin, & Anderson, 2010, p. 458).

#### 4.3 Classification

Classification models were evaluated for all participants and 17 subgroups by age, gender, and course. Model accuracy for six classification algorithms and two ensembles were compared. Three classification algorithms were linear classifiers, namely: Naïve Bayes (NB), Decision Tree (DT), and Logistic Regression (LR). Two were non-linear: Back Propagation Neural Network (BPNN) and k-Nearest Neighbour ( $k$ -NN). A Support Vector Machine (SVM) was trained both without (linear) and with (non-linear) a kernel function. A majority vote ensemble was trained using six base classifiers: two  $k$ -NN models trained on bootstrap samples of the data, SVM, NN, DT, and NB.<sup>9</sup> Finally, a bagging ensemble used the base classifier with the highest accuracy. RapidMiner version 5.3 (rapidminer.com) was used for modelling.

Two training methods were compared for models of all participants: 1) 10-fold cross validation using stratified sampling ( $\text{Model}_{xval}$ ); and 2) model accuracy when trained on 2010 and 2011 participants and tested on 2012 participants ( $\text{Model}_{2012}$ ). Subgroup models were trained on 2010 and 2011 participants and tested on 2012 participants.

---

<sup>9</sup> Including LR decreased model accuracy; using two  $k$ -NN learners increased model accuracy.

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

#### 4.3.1 Model parameter tuning

Parameter tuning for each algorithm, and the optimal parameter settings for models of all participants, were as follows:

- $k$ -NN models were trained on values of  $k$  in the range [2,50]. Best Model<sub>XVal</sub> used  $k=18$ ; best Model<sub>2012</sub> used  $k=15$ .
- SVM models were trained on three complexity constants ( $C$ ) in the range [0–5] and four kernel functions, namely dot (none), radial, polynomial, and anova. Best Model<sub>XVal</sub> used a radial kernel function and  $C=0$ ; best Model<sub>2012</sub> used a dot kernel function and  $C=0$ .
- Best accuracy for BPNN was achieved with a learning rate of 0.25, momentum of 0.3, 500 training cycles, and the default configuration of one hidden layer with number of neurons set to:  $\left(1 + \frac{\text{number of attributes} + \text{number of classes}}{2}\right)$ . The activation function was sigmoid; attributes were scaled to the range [-1,1].
- The decision tree required an information-gain threshold  $<0.05$  to train. Models were trained on information-gain thresholds in the range [0.0,0.05]; additional pruning was controlled by minimum leaf size in the range [2,20]. Best Model<sub>XVal</sub> used an information-gain threshold of 0.009 and a minimum leaf size of 10; best Model<sub>2012</sub> used an information-gain threshold of 0.016 and a minimum leaf size of 5.
- Bagging ensembles were tested with iteration values in the range [2,15] and sample ratios in the range [0.5,1]. Optimal bagging accuracy used 8 iterations and a sample ratio of 0.9, although all iteration values gave similar results. Parameter settings for the constituent base algorithm were as per configuration details above.
- The voting ensemble used majority voting. Parameter settings for the constituent base algorithms were as detailed above.

#### 4.3.2 Reporting model accuracy

Two results are reported for each model, accuracy and geometric mean (GM). GM is more appropriate than accuracy for unbalanced datasets; it combines the precision and recall of each class and so compensates for the greater influence of the majority class in accuracy calculations (Romero, Ventura, Espejo, & Hervás, 2008). As will be discussed in Section 5.4, the degree of class imbalance varied across subgroups in the dataset, which was addressed by over-sampling the minority class. Accuracy was calculated from the confusion matrix of the balanced dataset. GM was calculated from the confusion matrix of the original, unbalanced dataset, i.e., after removal of bootstrap replicas.

#### 4.3.3 Comparing classification model accuracies

Model accuracies were compared based on their confusion matrices. Two tests were used, McNemar's test and Fisher's exact test (FET). McNemar's test, based on chi squared ( $\chi^2$ ), can be used to compare the results of two classification models applied to the same dataset (Dietterich, 1998). When comparing several algorithms, p-values were adjusted using the Holm correction to account for family-wise error. FET was used to compare model accuracies when applied to different datasets; for example, comparing an algorithm's performance for Model<sub>XVal</sub> and Model<sub>2012</sub> (Rice, 1995, p. 484). Both tests were run in R.

#### 4.3.4 Attribute subset selection

Attribute subset selection techniques can improve model performance and identify relevant attributes

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

(Hall & Homes, 2003). Three wrapper methods were used for attribute subset selection, namely forward selection, backward selection, and a genetic algorithm. The two most popular approaches, forward selection and backward selection, generally give good results (Hall & Homes, 2003) but can converge on a local optimum (Baumann, 2003), hence a genetic algorithm (GA) was also used (Yang & Honavar, 1998). In an analysis of 144 optimal Model<sub>2012</sub> models across algorithms and dataset subgroups, forward selection identified the best model most frequently (65% of models), followed by Genetic algorithm (33% of models). Forward selection identified 82% of the best  $k$ -NN models, where best model referred to the model with the highest overall accuracy.

## 5 DATA PRE-PROCESSING

The following section discusses missing data, generating a binary class label from GPA, class imbalance, and assessing sample size. All attributes were scaled using a standard normal Z-transformation ( $m=0$ ,  $s=1$ ).

### 5.1 Missing Data on Prior Academic Performance

Prior academic performance was unavailable for 189 (16%) of the 1,207 study participants. An additional 20 students (2%) had fewer than the required six leaving certificate subjects, so CAO points underestimated prior academic performance. Both subgroups represented non-standard students of interest to the study and remained in the dataset. However, an additional factor, average leaving certificate results ( $m=38.4$ ,  $s=11.9$ ,  $n=1,018$ ), was added to indicate average results achieved across all leaving certificate subjects attempted.

### 5.2 Discretizing Academic Performance

The class label needed to identify students at risk of failing. Experimenting with various ways to discretize GPA, Minaei-Bidgoli, Kashy, Kortemeyer, and Punch (2003) found that higher numbers of bins (up to nine) resulted in higher errors because of low sample size in some bins ( $n=227$ ). They recommended using either two ( $GPA \leq 2.0$  and  $GPA > 2$ ) or three ( $GPA \leq 2$ ;  $2 < GPA < 3.5$ , and  $GPA \geq 3.5$ ) bins. Thirty percent of their participants achieved a GPA of 3.5 or higher. Romero et al. (2008) binned module grades into four categories based on final course mark (range [0,10]): fail [0,5], pass [5,7], good [7,9], and excellent [9,10]. However, they reported excellent and good students were frequently classified as pass, again citing lower sample sizes in those groups as a possible cause ( $n=438$ ). Thai-Nghe, Janecek, and Haddawy (2007) achieved best recall for failed students (64%) using two bins based on end-of-year GPA (range [2,4]), namely fail [2, 2.5) and pass [2.5, 4] ( $n=20,492$ ).

Both two and three GPA bins were considered for this study. To evaluate boundaries using two bins, seventeen GPA bin boundaries in the range [1.7, 2.5] were assessed using Naïve Bayes<sup>10</sup> with 10-fold

---

<sup>10</sup> Early models of the data suggested Naïve Bayes gave comparable accuracies to other learners, concurring with Bergin (2006), as discussed in Section 7.1.

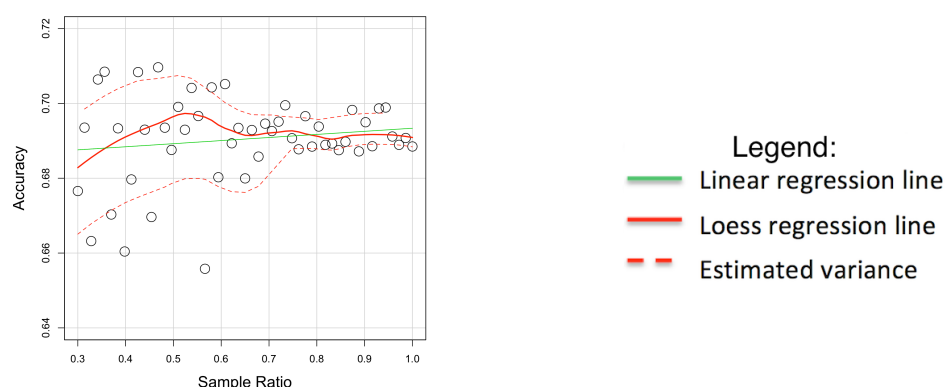
(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

cross validation. Optimal accuracy was achieved with a boundary of GPA=2.0 (accuracy: 68.5%, recall on fail: 71%) confirming a boundary between a passing and failing GPA. Models predicting three GPA bins were less successful. Sixteen models were tested using lower GPA boundary values in the range [1.0, 2.0] and upper GPA boundary values in the range [2.5, 3.25]. Models had difficulty distinguishing between medium- and low-risk students. The highest overall accuracy was achieved with GPA boundaries of 1.8 and 3.25 (accuracy: 53.5%, recall on fail: 64%), which was marginally better than a random guess ( $\kappa=0.3$ ).<sup>11</sup> Superby, Vandamme, and Meskens (2006) had similarly poor results predicting three classes. Therefore, two GPA bins were used for classification models in this study, GPA<2.0 (*class=fail*) and GPA≥2 (*class=pass*), distinguishing high-risk students from other students.

### 5.3 Evaluation of Sample Size

Progressive sampling indicates if combinations of attribute values likely to occur amongst study participants are sufficiently represented in a sample (Provost, Jensen, & Oates, 1999). Each classification algorithm was trained on fifty sample sizes between the sampling fractions of 0.3 ( $n=362$ ) and 1 ( $n=1,207$ ), using 10-fold cross validation. Variance in Naïve Bayes model accuracy converged for sample fractions >0.75 ( $n=905$ ). In addition, the slope of the LOESS regression line for Naïve Bayes model accuracy was approximately zero for sample fractions >0.8 ( $n=966$ ), indicating convergence of model accuracy as illustrated in Figure 1. Naïve Bayes can converge to optimal accuracy on a smaller sample size than other algorithms (Ng & Jordon, 2001; Mitchell, 2015). Variance in model accuracy for DT, BPNN, and  $k$ -NN converged for sample fractions >0.8 ( $n=966$ ). Variance in SVM model accuracy converged for sample fractions >0.85 ( $n=1,026$ ) and LR model accuracy appeared to converge for sample fractions >0.93 ( $n=1,123$ ), however a larger dataset would be needed to confirm accuracy convergence for LR.

**Key finding:** Sample sizes less than  $n=900$  under-represented patterns in the dataset.



**Figure 1: Model accuracy for progressive sampling using Naïve Bayes, generated using the scatterplot function in the R package *car*, version 2.0-21.**

<sup>11</sup> Cohen's Kappa coefficient ( $\kappa$ ), range -1 to 1 inclusive, is a measure of the extent to which this result could have occurred by chance;  $-0.2 < \kappa < 0.2$  indicates a performance similar to a random guess (Kundel & Polansky, 2003).



## 5.4 Class Imbalance

Thirty-eight percent of all participants were in class *fail*. The relative class size for *fail* varied across dataset subgroups, ranging from 15% to 56%. Results from progressive sampling indicated the dataset was too small to under-sample the majority class, therefore two class balance options were compared: basic over-sampling of the minority class and synthetic minority over-sampling technique (SMOTE). Basic over-sampling of the minority class has been criticized for not addressing the issue of lack of data, and for over-fitting the data (Weiss, 2004). Chawla, Bowyer, Hall, and Keelmeier (2002) proposed SMOTE as an alternative over-sampling approach that generates synthetic instances along line segments joining nearest neighbours in the minority class. Additional instances of the minority class were generated using the SMOTE function implemented in R (package DMwR version 0.4.1). Models were trained on the 2010 and 2011 student cohort and tested on the 2012 student cohort. Training and test datasets were balanced separately, i.e., instances from the training dataset were not available when re-sampling test instances, and vice versa. SMOTE resulted in lower model accuracies than simple over-sampling for all learners, therefore, simple over-sampling of the minority class was used. However, it is worth noting that a comparison of model accuracies using SMOTE versus simple over-sampling showed the differences were not statistically significant.

## 6 RESULTS

Results of statistical analysis exploring relationships between study factors and GPA are given to facilitate both comparison with other studies, and inform the discussion of classification model results in Section 7. This includes correlation analysis and analysis of group differences. Results from classification models predicting a binary class label of *fail* ( $GPA < 2.0$ ) and *pass* ( $GPA \geq 2.0$ ) are presented, including identification of key attributes used across classification models.

### 6.1 Correlations between Study Attributes

All factors of prior academic performance had significant correlations with each other and lower but significant correlations with GPA ( $p < 0.05$ ), illustrated as a heat map in Table 6. Methodical subjects ( $r = 0.302$ , 95% B-CI [0.24, 0.36]), CAO Points ( $r = 0.285$ , 95% B-CI [0.22, 0.34]), and Mathematics ( $r = 0.274$ , 95% B-CI [0.21, 0.33]) had the highest correlations with GPA. Similar correlations were cited in other studies that included mature students (Conrad, 2006; Duff, Boyle, Dunleavy, & Ferguson, 2004; Kaufman, Agars, & Lopez-Wagner, 2008).

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372.  
<http://dx.doi.org/10.18608/jla.2016.32.20>

**Table 6: Correlations between factors of prior academic performance and GPA.**

	GPA	CAO Points	English	Mathematics	Applied Average	Humanities Average
CAO Points	0.285 [0.22, 0.34]					
English	0.169 [0.11, 0.23]	0.698 [0.67, 0.73]				
Mathematics	0.274 [0.21, 0.33]	0.477 [0.42, 0.53]	0.251 [0.19, 0.31]			
Applied Average	0.172 [0.10, 0.24]	0.560 [0.50, 0.61]	0.365 [0.29, 0.43]	0.173 [0.09, 0.25]		
Humanities Average	0.228 [0.17, 0.29]	0.820 [0.79, 0.84]	0.693 [0.66, 0.73]	0.263 [0.20, 0.32]	0.338 [0.26, 0.40]	
Methodical Average	0.302 [0.24, 0.36]	0.707 [0.68, 0.74]	0.432 [0.38, 0.48]	0.681 [0.65, 0.71]	0.194 [0.12, 0.26]	0.418 [0.36, 0.47]

Intervals are 95% Confidence Intervals based on 1,999 bootstrap samples. Only students with school leaving certificate results were included (n=1,018). Applied average results are based on a subset of students who did applied subjects (n=647, 64%).

With the exception of visual and auditory modality, all non-cognitive factors of learning were significantly correlated with GPA ( $p < 0.05$ ). Table 7 is a heat map visualization of correlations for non-cognitive study factors; correlations with confidence intervals are given in the Appendix, Table 15. *Age* ( $r = 0.25$ , 95% B-CI [0.2, 0.3]), a *deep learning approach* ( $r = 0.234$ , 95% B-CI [0.18, 0.29]), and *study effort* ( $r = 0.187$ , 95% B-CI [0.14, 0.24]) had the highest correlations with GPA. *Openness* ( $r = 0.084$ , 95% B-CI [0.03, 0.14]) and *group work* ( $r = -0.08$ , 95% B-CI [-0.13, -0.02]) had the weakest statistically significant correlations with GPA. Correlations were comparable with other studies of diverse student populations with the exception of *self-efficacy* ( $r = 0.12$ , 95% B-CI [0.06, 0.18]), which was lower than expected (Cassidy, 2011:  $r = 0.397$ ; Diseth, 2011:  $r = 0.44$ ; Komarraju & Nadler, 2013:  $r = 0.30$ ).

The relatively low validity and internal reliability of *intrinsic goal orientation* was not reflected in correlations with other attributes. All factors of motivation were statistically significantly correlated with each other. The highest correlation was between *intrinsic goal orientation* and *self-efficacy* ( $r = 0.421$ , 95% B-CI [0.37, 0.47]), which concurred with Diseth (2011) and was marginally lower than Komarraju and Nadler (2013). Also of note was the statistically significant correlation between *intrinsic* and *extrinsic goal orientation* ( $r = 0.381$ , 95% B-CI [0.33, 0.43]) as correlations cited in other studies were inconsistent (Diseth, 2011; Eppler & Harju, 1997; Komarraju & Nadler, 2013). Correlations between factors of motivation and factors of self-regulation were also statistically significant, particularly *intrinsic goal orientation* and a *deep learning approach* ( $r = 0.417$ , 95% B-CI [0.37, 0.47]).

*Study time* had relatively low validity and internal reliability as discussed in Section 3.2.2. While correlations with *metacognitive self-regulation* (0.452, 95% B-CI [0.4, 0.49]) and *study effort* ( $r = 0.378$ , 95% B-CI [0.33, 0.43]) were statistically significant, they were lower than results cited in other studies; for example, Bidjerano and Dai (2007) ( $r = 0.55$  and  $r = 0.64$  respectively), which was based on a similar

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

participant profile.

The high negative correlation between a *deep* and *strategic learning approach* ( $r=-0.791$ , 95% B-CI [-0.81, -0.77]) reflected that most participants reported they were not shallow learners ( $m=1.3$ ,  $s=1.9$ ,  $range=[0,10]$ ), selecting either deep or strategic statements. As expected, a *shallow learning approach* was negatively correlated with other non-cognitive factors. However, a *strategic learning approach* was also negatively correlated with other factors of learning and GPA ( $r=-0.158$ , 95% B-CI [-0.22, -0.10]), contradicting other studies (e.g., Duff et al., 2004; Swanberg & Martinsen, 2010). The difference may be explained by their questionnaire design, which facilitated selection of both *strategic* and *deep learning approaches*, resulting in a statistically significant positive correlation between *deep* and *strategic learning approaches*.

**Key finding:** With the exception of *study time*, correlations between study factors concurred with results cited in other studies. Also of note was the relatively low correlation between *self-efficacy* and GPA.

## 6.2 Group Differences

### 6.2.1 GPA bands

Group differences were assessed for the three GPA bands discussed Section 3.2.3, namely high-risk ( $GPA < 2.0$ ), medium-risk ( $2.0 \leq GPA < 2.5$ ), and low-risk ( $GPA \geq 2.5$ ) students. A *deep learning approach* was the only attribute with statistically significant differences across the three groups ( $F(2,1204)=25.95$ ,  $p < 0.001$ ). High-risk students had significantly lower prior academic performance than either medium- or low-risk students, particularly in *methodical average* ( $F(2,1015)=59.98$ ,  $p < 0.001$ ), *CAO points* ( $F(2,1015)=50.33$ ,  $p < 0.001$ ), and *mathematics* ( $F(2,1015)=46.02$ ,  $p < 0.001$ ). Low-risk students were significantly different from the other two groups in some effective learning dispositions. They had higher scores in *intrinsic goal orientation* ( $F(2,1015)=50.22$ ,  $p < 0.001$ ), *study effort* ( $F(2,1204)=17.76$ ,  $p < 0.001$ ), *conscientiousness* ( $F(2,1204)=11.42$ ,  $p < 0.001$ ), and *openness* ( $F(2,1204)=5.77$ ,  $p < 0.003$ ). For the remaining non-cognitive factors, there were statistically significant differences between high- and low-risk students only, but medium-risk students did not differ significantly from the other two groups..

**Table 7: Correlations for non-cognitive factors.**

SE	Personality			Motivation			Self-regulation			Learning Approach			Other			Modality	
	GPA	Con	Open	SE	EM	IM	SR	StE	StT	Deep	Stra	Shal	Group	Age	Gen	Vis	Aud
Con	0.150			0	0	0	0	0	0	0	0	0	0	0	0	0	0
Open	0.084	0.032			0	0	0	0	0	0	0	0	0	0	0	0	0
SE	0.120	0.313	0.178		0	0	0	0	0	0	0	0	0	0	0	0	0
EM	0.124	0.280	0.049	0.308			0	0	0	0	0	0	0	0	0	0	0
IM	0.149	0.334	0.316	0.421	0.381		0	0	0	0	0	0	0	0	0	0	0
SR	0.130	0.515	0.101	0.409	0.298	0.429		0	0	0	0	0	0	0	0	0	0
StE	0.187	0.450	0.064	0.334	0.232	0.330	0.594		0	0	0	0	0	0	0	0	0
StT	0.101	0.396	0.009	0.259	0.175	0.227	0.452	0.378		0	0	0	0	0	0	0	0
Deep	0.234	0.352	0.209	0.273	0.158	0.417	0.431	0.360	0.285			0	0	0	0	0	0
Stra	-0.158	-0.167	-0.174	-0.158	-0.012	-0.274	-0.213	-0.133	-0.115	-0.791			0	0	0	0	0
Shal	-0.146	-0.330	-0.096	-0.221	-0.234	-0.294	-0.398	-0.394	-0.290	-0.519	-0.103		0	0	0	0	0
Group	-0.080	0.052	-0.042	0.056	0.059	0.027	0.113	0.094	0.084	0.020	0.037	-0.081			0	0	0
Age	0.250	0.156	0.038	0.038	0.051	0.257	0.234	0.210	0.023	0.284	-0.200	-0.181	-0.022			0	0
Gen	0.100	-0.005	0.022	-0.048	0.035	0.004	0.005	0.023	0.086	0.086	-0.001	-0.130	0.026	-0.038		0	0
Vis	0.050	0.069	0.063	-0.024	0.041	0.054	0.024	-0.003	0.038	0.067	-0.020	-0.089	0.021	-0.038	-0.046		0
Aud	0.020	0.073	0.023	-0.002	0.013	-0.016	0.065	0.039	0.081	0.077	-0.068	-0.026	-0.097	0.025	0.205	-0.347	
Kin	-0.059	-0.124	-0.074	0.022	-0.046	-0.032	-0.078	-0.033	-0.105	-0.126	0.078	0.099	0.069	-0.055	-0.144	-0.541	0.601

Con=Conscientiousness; Open=Openness; SE=Self-efficacy; IM=Intrinsic goal orientation; EM=Extrinsic goal orientation; SR=Metacognitive self-regulation; StE=Study effort; StT=Study time; Deep=Deep Learner; Shal=Shallow learner; Stra=Strategic learner; Group=Likes to work in groups; Gen=Gender; Vis=Visual learner; Aud=Auditory learner; Kin=Kinaesthetic learner.

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

Groups did not differ significantly in *visual* or *auditory modality*. Group means are given in the Appendix, Table 13.

**Key finding:** Group differences by GPA band confirmed that both prior academic performance and non-cognitive factors associated with an effective learning disposition differentiated high-risk from low-risk students

### 6.2.2 Age categories

Study factor mean and standard deviations were initially compared for ten age categories<sup>12</sup>: 18, 19, 20, 21, 22–23, 24–25, 26–28, 29–32, 33–39, and ≥40. Age groups were combined to ensure at least 60 students per group.<sup>13</sup> Comparing differences in means reduced the ten categories to three, namely: 18–23 (n=875); 24–28 (n=131); and 29–60 (n=201). This was based on a lack of significant differences for study factors in age groups within these three age categories. Analysis of differences across the three age categories showed average GPA scores increased significantly with *age* ( $F(2,1204)=48.95$ ,  $p<0.001$ ), as did many non-cognitive factors associated with an effective learning disposition, namely a *deep learning approach* ( $F(2,1204)=68.54$ ,  $p<0.001$ ), *intrinsic goal orientation* ( $F(2,1204)=51.6$ ,  $p<0.001$ ), *metacognitive self-regulation* ( $F(2,1204)=39.19$ ,  $p<0.001$ ), *study effort* ( $F(2,1204)=32.57$ ,  $p<0.001$ ), *conscientiousness* ( $F(2,1204)=16.06$ ,  $p<0.001$ ), and *extrinsic goal orientation* ( $F(2, 1204)=5.287$ ,  $p<0.01$ ). As expected, *CAO points* decreased with *age* ( $F(2,1015)=54.08$ ,  $p<0.001$ ) as entry requirements are lower for students aged 23 and over. This difference in prior academic performance was reflected in all subject areas except *mathematics* ( $F(2, 1015)=0.271$ ,  $p=0.763$ ). These results concur with a number of studies reporting a better learning disposition and academic performance amongst older students (e.g., Cassidy, 2011; Eppler & Harju, 1997; Hoskins, Newstead, & Dennis, 1997). Group means are included in the Appendix, Table 13.

**Key finding:** Group differences by *age* concurred with other studies that older students had a more effective learning disposition and higher GPA.

### 6.2.3 Gender differences

Engineering and computing courses were predominantly male and had low entry requirements. Humanities courses were predominantly female and had high entry requirements. Therefore, it is unsurprising that males had significantly lower *CAO points* than females ( $t(918)=-4.077$ ,  $p<0.001$ ). This difference was reflected in all subject areas, the least significant difference was in *mathematics* ( $t(913)=-2.081$ ,  $p=0.038$ ), as illustrated in Table 13 in the Appendix. GPA scores were also significantly lower for males ( $t(1158)=-3.595$ ,  $p<0.001$ ). Gender differences in academic performance were not reflected in factors of temperament or motivation, but statistically significant differences were found in factors of self-regulation, approaches to learning, and learner modality. Females had higher mean scores for *study time* ( $t(1065)=-2.988$ ,  $p=0.003$ ), a *deep learning approach* ( $t(1107)=-3.038$ ,  $p=0.002$ ), and *auditory modality* ( $t(1170)=-7.300$ ,  $p<0.001$ ) while males had higher mean

<sup>12</sup> Student age was calculated as whole years old at September 1<sup>st</sup> in the year of registration.

<sup>13</sup> Applying both a Wilcoxon Rank Sum non-parametric test and a parametric t-test to a range of distributions common in psychometric data, Micceri (1989) reported t-tests inflated Type I errors (incorrect finding of significance) for sample sizes less than 60 only, but performed well for larger samples.



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

scores for a *shallow learning approach* ( $t(1170)=4.723$ ,  $p<0.001$ ) and *kinaesthetic learning modality* ( $t(1131)=5.175$ ,  $p<0.001$ ).

## 6.3 Classification Results

### 6.3.1 Classification accuracy for models of all participants

Table 8 gives classification model accuracies for models of all participants as a heat map. As discussed in Section 5.4, the dataset was balanced by oversampling the minority class. Reported model accuracies are based on the resulting balanced dataset. However, geometric mean (GM) was calculated after removal of replicated instances from the labelled dataset as explained in Section 4.2.2. This is indicated by the sample size ( $n$ ) included in Table 8. Contingency tables for McNemar's test ( $\chi^2$ ) and Fisher's exact test (FET) also excluded replicated instances.

Best Model<sub>2012</sub> accuracy was  $k$ -NN (accuracy: 71.98%, GM: 70.35%). However, a comparison of Model<sub>2012</sub> accuracies using McNemar's test with Holm correction for family-wise error indicated model performance was comparable across algorithms. The only statistically significant difference was between LR (accuracy: 65.38%, GM: 61.12%) and  $k$ -NN ( $\chi^2(1, n=436)=15.95$ ,  $p<0.001$ ).

In contrast with Model<sub>2012</sub> results, best Model<sub>XVal</sub> accuracy was SVM (81.62%). Its lower GM (72.18%) reflected a higher recall on *pass* (88.24%) than *fail* (59.04%). Given the objective of identifying students at risk of failing, models with a high recall on *fail* are preferable. BPNN also had good accuracy (75.33%) but a lower GM (69.32%) reflecting a higher precision on *pass* (78.53%) than *fail* (57.48%). A comparison of Model<sub>XVal</sub> accuracies using McNemar's test with Holm correction showed SVM had statistically significantly higher accuracy than all other algorithms. In addition, LR accuracy was statistically significantly lower than other algorithms (accuracy: 66.64%, GM: 63.06%). Differences in the accuracies of the remaining six algorithms were not statistically significant.

Model<sub>XVal</sub> accuracy was higher than Model<sub>2012</sub> accuracy for each algorithm used. However, a comparison of the Model<sub>2012</sub> and Model<sub>XVal</sub> accuracy for each algorithm found a statistically significant difference for SVM only ( $p<0.001$  FET).<sup>14</sup>

Extensive search strategies were used for optimal attribute subset selection as discussed in Section 4.3.4. This can result in model over-fitting (Baumann, 2003). A comparison of cross validation accuracies with (Model<sub>XVal</sub>), and without (Model<sub>all</sub>), attribute subset selection found that the difference was statistically significant for two algorithms only,  $k$ -NN ( $\chi^2(1, n=1207)=20.1$ ,  $p<0.001$ ) and a  $k$ -NN Bagging Ensemble ( $\chi^2(1, n=1207)=6.7$ ,  $p<0.01$ ). Model<sub>all</sub> accuracies are included in Table 8.

**Key finding:** Models of students at risk of failing based on factors measured prior to commencement of first year of study achieved good predictive accuracy. Model accuracies were comparable across a number of classifiers.

<sup>14</sup> Fisher's exact test (FET) assumes independent samples; it is likely that some instances from Model<sub>XVal</sub> were replicated in Model<sub>2012</sub>, violating the assumptions of FET. Significance concurred with results from a selection of other statistical tests including chi-squared and Z-score probability.

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

### 6.3.2 Optimal attribute subsets used

Table 9 illustrates optimal attribute subsets used by each of the eight  $Model_{2012}$  classification algorithms when modelling all participants. The five factors most frequently used were *age*, *methodical average*, *leaving certificate average*, *self-efficacy*, and *kinaesthetic modality*. *English* and *auditory modality* were ignored by all algorithms while *CAO points*, *mathematics*, *conscientiousness*, and *study time* were each selected by just one algorithm.

**Table 8: Heat map of classification model accuracies for all participants.**

Algorithm	$Model_{2012}$		$Model_{XVal}$		$Model_{All}$
	Accuracy (%) (n=546)	GM (%) (n=436)	Accuracy (%) (n=1,496)	GM (%) (n=1,207)	Accuracy (%) (n=1,496)
k-NN	71.98	70.35	72.39	71.38	66.12
Voting	71.06	68.89	74.26	68.86	72.73
Bagging	70.51	68.78	71.93	70.51	67.45
SVM	70.33	67.32	81.62	72.18	81.48
DT	69.96	68.38	70.19	69.32	69.59
NB	66.30	64.73	69.52	69.14	68.78
BPNN	68.50	66.29	75.33	68.97	76.40
LR	65.38	61.12	66.64	63.06	64.91

GM=Geometric mean.

**Table 9: Attributes used by  $Model_{2012}$  models.**

Attribute	SVM	Voting	BPNN	k-NN	DT	LR	Bagging	NB	Count
Age	*	*	*	*	*	*	*		7
Methodical average	*	*	*	*			*	*	6
Leaving certificate avg.	*	*	*		*	*	*		6
Self-efficacy	*	*		*	*	*	*		6
Kinaesthetic	*	*	*	*	*	*			6
Humanities average	*	*	*	*		*			5
Intrinsic goal orientation	*	*	*		*	*			5
Openness	*	*	*		*				4
Gender	*	*	*			*			4
Deep learner		*	*			*			3
Applied average				*			*		2
Extrinsic goal orientation	*							*	2
Metacognitive self-regulation	*					*			2
Study effort		*		*					2
Shallow learner	*		*						2
Group work	*						*		2
Visual				*	*				2
Strategic learner		*	*						2
Mathematics					*				1
CAO Points				*					1
Conscientiousness	*								1
Study time					*				1

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

Attribute	SVM	Voting	BPNN	k-NN	DT	LR	Bagging	NB	Count
Auditory									0
English									0
Number of attributes used:	14	12	11	9	9	9	6	2	

\*Attribute was included in the model; dashed lines were added to improve readability.

To compare the predictive accuracy of non-cognitive factors of learning with the predictive accuracy of factors available from student registration, three models were compared: a  $k$ -NN model trained on all attributes ( $Model_{2012}$ ), a  $k$ -NN model trained on prior academic performance, age, and gender ( $Model_{prior}$ ), and a  $k$ -NN model trained on the non-cognitive learning factors of age and gender ( $Model_{NCog}$ ). Models were trained on 2010 and 2011 data, and then tested on the 2012 data. Forward selection was used for attribute subset selection and models were trained on values of  $k$  in the range [10,30]. Accuracies were compared using McNemar's test with Holm correction.  $Model_{prior}$  accuracy (70.33%) was marginally lower than  $Model_{2012}$  accuracy (71.88%), but the difference was not statistically significant ( $p=0.44$ ).  $Model_{NCog}$  accuracy (64.10%) was statistically significantly lower than  $Model_{2012}$  ( $p=0.04$ ) but was not statistically significantly lower than  $Model_{prior}$  ( $p=0.12$ ). A subset of four attributes was used in  $Model_{prior}$ , namely *average leaving certificate*, *methodical average*, *age*, and *gender*. A subset of nine attributes was used in  $Model_{NCog}$ , namely *conscientiousness*, *self-efficacy*, *intrinsic and extrinsic goal orientation*, *study effort*, *deep and shallow learning approaches*, *age*, and *gender*.

**Key findings:** The five attributes most predictive of students at risk of failing were *age*, *methodical average*, *leaving cert average*, *self-efficacy*, and *kinaesthetic modality*. Improvement in model accuracy attributed to non-cognitive factors of learning was not statistically significant.

**Table 10: Confusion matrix.**

	Predicted Fail	Predicted Pass	Recall
Actual Fail	116 ( <i>True Fail</i> (TF))	47 ( <i>False Pass</i> (FP))	71.17% $\frac{TF}{TF + FP}$
Actual Pass	87 ( <i>False Fail</i> (FF))	186 ( <i>True Pass</i> (TP))	68.13% $\frac{TP}{FF + TP}$
Precision	57.14% $\frac{TF}{TF + FF}$	79.83% $\frac{TP}{TP + FP}$	69.27% $\frac{TF + TP}{TF + TP + FF + FP}$

### 6.3.3 Group differences between misclassifications and correct predictions

A review of  $Model_{2012}$  predictions highlighted that algorithms generally concurred on participants misclassified. For example, 75% of participants misclassified by  $k$ -NN were misclassified by at least four of the eight algorithms used. Therefore, student misclassification was defined as an instance misclassified by at least four of the eight algorithms. The resulting confusion matrix identified four groups as illustrated in Table 10: students correctly predicted as fail (*True Fail*); students correctly predicted as pass (*True Pass*); students incorrectly predicted as fail (*False Fail*); and students incorrectly predicted as pass (*False Pass*). A Shapiro-Wilk test of fifty bootstrap samples of each group verified group means were normally distributed for each study factor but a Levene's test found variances were unequal. Therefore, Welch's t-test was used to compare each misclassified

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

group with correctly classified participants. Results and group means are given in Table 11. Group differences concurred with a similar analysis of *k*-NN misclassifications.

Of specific interest was each misclassified group and how it differed from correctly classified participants. *False Pass* had similar mean scores to *True Pass* in all study attributes. However, *False Pass* had statistically significantly higher mean scores than *True Fail* in a number of non-cognitive factors of learning, *methodical average*, and *mathematics*. *False Fail* had a statistically significantly lower GPA than *True Pass*, and had similar mean scores to *True Fail* in all study factors except *CAO points*, *humanities average*, and *English*.

**Key finding:** Analysis of group differences between instances incorrectly predicted as pass and instances correctly predicted as pass showed a lack of statistically significant differences between the two groups. Conversely, differences between instances incorrectly predicted as pass and instances correctly predicted as fail were statistically significant in a number of effective learning dispositions and measures of prior academic performance

#### 6.3.4 Classification models by subgroup

Classification models were trained for subgroups by academic course (12 subgroups), gender, and age (3 subgroups). As with the full dataset, eight classification algorithms were trained on 2010 and 2011 data and tested on 2012 data. Subgroup model accuracies were higher than models trained on the full dataset. However, all subgroups had  $n < 900$  and so were below the minimum sample size required to accurately model study factors as discussed in Section 5.4. Therefore, to assess if insufficient data contributed to high model accuracies, model accuracy for each subgroup was compared to the mean accuracy of models trained on 50 random bootstrap samples of the same size, selected from all participants. The random samples were constructed to match the class imbalance of their corresponding subgroup, i.e., the numbers of passes and fails matched for both test datasets and training datasets. Average model accuracies for random samples were also higher than models of all participants, and in many cases were similar to their corresponding subgroup model accuracy. This suggested that under sampling of the data contributed to improvement in model accuracy, invalidating results from models of subgroup.

**Key finding:** Sample sizes were too small to draw conclusions on subgroup model accuracies.

## 7 ANALYSIS OF RESULTS

Results from this study indicated that predictive models of students failing in their first year of study in tertiary education, trained on data available during first-year student induction, can achieve good predictive accuracy when applied to a different student cohort. Therefore, the study corroborates the use of data mining's empirical modelling approaches for predicting at-risk students. The following sections discuss the salient outcomes from this study.

### 7.1 Models of Academic Performance

The study dataset was diverse in terms of student age and course of study. Modes of assessment also varied across courses. For example, humanities courses gave more weighting to end-of-term examinations, other courses gave equal or higher weighting to continuous assessment work.

**Table 11. Group differences for misclassified participants, including group means ( $m \pm s$ ).**

Attribute	True Fail n=116	False Fail n=87	False Pass n=47	True Pass n=186	False Pass compared with:		False Fail compared with:	
					True Fail	True Pass	True Fail	True Pass
CAO points	207.80±66.27	235.98±60.94	236.4±130.92	236.27±141.43			**	
Mathematics	16.45±11.51	17.76±12.08	23.00±16.67	23.72±17.28	*			**
English	40.66±18.81	45.81±15.19	37.34±25.64	42.63±27.09			*	
Methodical average	21.41±10.50	24.28±10.59	33.15±19.61	32.97±20.96	***			***
Creative average	31.69±26.67	32.98±27.95	25.53±26.65	22.31±29.39				**
Humanities average	32.77±12.49	37.22±11.55	36.45±22.00	35.94±23.01			**	
Conscientiousness	5.89±1.68	5.90±1.45	6.38±1.55	6.47±1.42				**
Openness	5.94±1.42	5.82±1.13	6.14±1.37	6.20±1.35				*
Self-efficacy	6.72±1.35	6.83±1.39	7.32±1.30	7.27±1.27	**			*
Intrinsic goal orientation	6.76±1.46	6.70±1.33	7.45±1.26	7.32±1.25	**			***
Extrinsic goal orientation	7.46±1.50	7.79±1.24	8.01±1.23	8.01±1.15	*			
Metacognitive SR	5.66±1.51	5.55±1.18	6.28±1.33	6.24±1.31	*			***
Study effort	5.39±1.94	5.65±1.64	6.35±1.54	6.26±1.74	**			**
Study time	6.06±2.49	6.04±2.07	6.14±2.27	6.53±2.18				
Deep learner	4.78±2.61	5.06±2.61	6.22±2.32	6.22±2.96	***			**
Strategic learner	3.73±2.43	3.79±2.44	2.71±2.07	2.81±2.51	**			**
Shallow learner	1.49±1.84	1.15±1.74	1.06±1.54	0.97±1.58				
Age	20.06±2.12	20.40±2.14	24.34±8.05	24.73±8.41	***			***
Group work	7.09±3.35	6.32±3.23	6.81±3.16	6.60±3.29				
Visual	7.13±2.04	7.27±2.14	7.18±1.78	7.16±2.08				
Auditory	3.06±2.11	3.36±2.15	3.14±2.24	3.07±2.15				
Kinaesthetic	4.81±2.38	4.37±2.70	4.68±2.31	4.77±2.49				
GPA	0.83±0.70	2.64±0.37	1.03±0.76	2.83±0.41		***	***	***

SR=self-regulation; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

Assessment methods can affect academic performance (Pérez-Martínez, García-García, Perdomo, & Villamide-Díaz, 2009) and its relationships with factors such as openness and learning approach, as discussed in Section 2. Notwithstanding these sources of variability, classification model accuracy was high (71.98%) when applied to a different student cohort.

A number of studies modelling educational data have cited comparable accuracies between classification algorithms, although there are inconsistencies regarding which model achieves optimal accuracy when modelling educational data. For example, Jayaprakash et al. (2014) reported SVM, LR, and NB had comparable accuracy predicting high-risk students based on factors including SAT scores, enrollment data, and data from an online learning environment, however DT did not perform as well ( $n=15,150$ ). On the other hand, Lauria, Moody, Jayaprakash, Jonnalagadda, and Baron (2013) reported DT had comparable accuracy with both SVM and LR when distinguishing between strong and weak students, also based on prior academic performance, demographic data, and log data from an intelligent tutoring system ( $n=6,445$ ). Herzog (2006) found DT and BPNN had similar performance to LR provided that independent variables had little co-linearity, but LR had lower accuracy when variables with greater dependencies were included in the model ( $n=4,564$ ). Bergin (2006) found NB and a stacking ensemble outperformed DT, BPNN,  $k$ -NN ( $k=3$ ), and LR when classifying students as strong or weak based on prior academic performance and psychometric data ( $n=102$ ). Results from this study concurred that a range of classification algorithms achieved similar accuracy; the highest accuracies were  $k$ -NN, ensembles, and SVM, as illustrated in Table 8.

While model accuracy estimated using 10-fold cross validation ( $\text{Model}_{\text{xVal}}$ ) was higher than model accuracy when tested on a different student cohort ( $\text{Model}_{2012}$ ), the increase was statistically significant for one algorithm only, SVM. As discussed in Section 4.3, a kernel function improved SVM model accuracy for  $\text{Model}_{\text{xVal}}$ , generating a more complex model than  $\text{Model}_{2012}$  where a kernel function failed to improve model accuracy. In contrast, optimal parameter settings for both  $\text{Model}_{2012}$  and  $\text{Model}_{\text{xVal}}$  were similar for other algorithms; for example,  $k$  in  $k$ -NN (15 and 18 respectively). Therefore, results indicated that cross validation provided a good estimate of model accuracy with the exception of a non-linear SVM model.

## 7.2 Analysis of Misclassifications

The misclassifications of particular interest were participants incorrectly predicted as *Pass*. Group comparisons of study attributes failed to identify differences between this group and those correctly predicted as *Pass*. There may be a number of reasons for this. First, a number of factors relevant to retention and progression arise after student induction, such as academic and social integration, change in circumstance resulting in economic pressure (Tinto, 2006), and classroom related affects on academic performance such as teaching methods (Ganyaupfu, 2013; Hake, 1998). Such factors may explain why prior academic performance and learning disposition alone are insufficient to predict academic performance in all cases. Second, it could be argued that profiling learners during first-year student induction is too early in the semester to measure some study attributes accurately. For example,

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

*intrinsic* and *extrinsic goal orientation* may vary depending on the time or situation (Apter, 1989). Similarly, students may be unsure of study expectations during the initial period of induction; Winters, Greene, and Costich (2008) concluded that both learner and task characteristics influenced levels and methods of self-regulation. On the other hand, factor correlations in this study (see Section 6.1) concurred with evidence cited in other studies where data were gathered later in the semester. Therefore, it was more likely that factors not included in the study explained incorrect predictions of *Pass*. Further work is needed to determine the potential improvements in model accuracy if additional data gathered after student induction were included.

Analysis of participants incorrectly predicted as *fail* showed they had a lower GPA than students correctly predicted as *pass*. Scores in a range of effective learning dispositions and prior academic performance were also lower, characterizing a group of lower academic achievers who may benefit from additional support in developing an effective learning disposition.

### 7.3 Overview of Study Attributes

The recorded significant correlations between study attributes meant different attribute subsets achieved comparable accuracies when predicting GPA. Consequently, no common subset of attributes could be isolated for use in this project. For example, a *deep learning approach* is associated with *intrinsic learning goals* (Gray, McGuinness, Owende, & Carthy, 2014) and correlations between the two factors were relatively strong ( $r=0.417$ , 95% B-CI [0.37, 0.47]). Six of the eight models used either a *deep learning approach* or *intrinsic goal orientation*, but just two models, BPNN and LR, used both. *k*-NN used neither of these, but instead used *self-efficacy*, which had a relatively high correlation with *intrinsic goal orientation* ( $r=0.421$ , 95% B-CI [0.37, 0.47]), and *study effort*, which had a relatively high correlation with a *deep learning approach* ( $r=0.360$ , 95% B-CI [0.31, 0.41]). Nevertheless, it is clear from reviewing attribute subsets used (Table 9) that some attributes were more predictive of students at risk of failing than others. The following sections provide an overview of the study factors and their relative usefulness in predicting students at risk of failing.

#### 7.3.1 Age and Gender

Correlations between *age* and academic performance are well cited in literature (Hoskins et al., 1997; Cassidy, 2011; Wigfield, Eccles, & Pintrich, 1996), and evidence from this study concurs with this observation. It is evident from analysis of group differences discussed in Section 6.2.2 that older students have a more effective learning disposition: they were more likely to adopt a deeper learning approach, set learning goals, and regulate learning. Classification models concurred that *age* was a good predictor of academic performance (see Table 9).

A number of studies reported *gender* is not significant in predicting academic performance in tertiary education (Naderi, Abdullah, Sharir, & Kumar, 2009; Dollinger, Matyja, & Huber, 2008; Hoskins et al., 1997); however, four of the eight algorithms used *gender* in spite of its relatively low correlation with GPA ( $r=0.1$ , 95% B-CI [0.05, 0.15]). Gender group differences highlighted that males had lower prior

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

academic performance, lower GPA, and lower scores in a number of effective learning dispositions. However, the dataset contained bias: courses that were predominantly male had lower entry requirements than courses that were predominantly female. Therefore, further work is needed to assess if the significance of *gender* in classification models was reflective of their course of study rather than actual gender differences.

### 7.3.2 Prior academic performance

Aggregate scores of prior academic performance, particularly *methodical average* and overall average (*leaving cert average*), were found to be more predictive of first-year academic performance than individual grades in *English* or *mathematics*. *Methodical average* had the highest correlation with GPA, and was used by most classification models. Replacing *methodical average* with *mathematics* reduced model accuracy. For example, *k*-NN model accuracy dropped from 71.3% to 66.29% ( $t(995)=39.1$ ,  $p<0.001$ ) when *methodical average* was replaced by *mathematics*. Therefore, while both factors displayed comparable correlations with GPA, an aggregate covering result in mathematics, science, and business subjects achieved higher accuracy when predicting students at risk of failing than *mathematics* alone.

A number of classification models included the factor *humanities average*, particularly models with higher accuracies such as *k*-NN, voting ensemble, SVM, and BPNN. Analysis of correlations with GPA by course of study showed that *humanities average* was more predictive of GPA in business and humanities courses than courses in engineering and computing (Appendix, Table 16). Correlations between leaving certificate *English* and GPA also varied by discipline. *English* had a statistically significant positive correlation with GPA for humanities courses and some business courses, but had a statistically significant negative correlation with GPA for Computing (IT). Correlations between *English* and GPA for other technical disciplines were not statistically significant ( $p>0.05$ ). Therefore, it was unsurprising that *English* was not a significant factor in classification or regression models of all participants.

### 7.3.3 Factors of personality

As discussed in Section 2, *conscientiousness* is the best personality-based predictor of academic performance in tertiary education, particularly for younger students (Kappe & van der Flier, 2010; Chamorro-Premuzic & Furnham, 2008; Kaufman et al., 2008; Allick & Realo, 1997). While correlation results from this study concurred with the available evidence, all classification models except SVM ignored *conscientiousness*. This suggested that other study factors accounted for *conscientiousness*, and there is no additional predictive value in measuring *conscientiousness* specifically.

*Openness* was used by four classification models (*k*-NN, BPNN, SVM, and voting), suggesting it was a useful predictor of students at risk of failing as evidenced by statistically significant group differences in *openness* for students with GPA  $>2.5$ . The mix of assessment methods used across courses and within courses may explain the low correlation between *openness* and GPA. Openness is the most controversial of the BIG-5 personality factors in terms of defining both meaning and sub-factors (de Raad & Schouwenburg, 1996, p. 321). The six-question scale used in this study covered four sub-factors:

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

creativity (2 questions), intellect (2 questions), imagination, and openness to new experiences (1 question each). Creativity, specifically, is frequently cited as an effective learning disposition that is to be encouraged and promoted in assessment design (Buckingham Shum & Deakin Crick, 2012). Further work is required to investigate if sub-factors inherent in openness may be more appropriate predictors of academic performance than openness itself.

#### 7.3.4 Factors of motivation

Results from this study support findings by Robbins et al. (2004) regarding the importance of *self-efficacy*; it was the most predictive non-cognitive factor of students at risk of failing. Interestingly, Deakin Crick and Goldspink (2014) observed that being able to express confidence in learning ability was also a strong indicator of an effective learning disposition. *Intrinsic goal orientation* was also indicative of good academic performance, however inferences on its relative importance must consider the poor reliability for that factor, as discussed in Section 3.2.2.

#### 7.3.5 Factors of self-regulation

Classification models largely ignored factors of self-regulation. The importance of self-regulated learning is well cited (e.g., Zimmerman, 1990); however self-regulation is complex to define and is related to a number of other factors of learning (Roll & Winne, 2015). For example, in a longitudinal study on the causal dilemma between motivation and self-regulation, De Clercq, Galand, and Frenay (2013) concluded that a learning goal orientation resulted in a deep learning approach, which in turn resulted in better self-regulation. Self-regulation, as measured in this study, failed to improve model performance over and above factors of motivation and approaches to learning.

#### 7.3.6 Approaches to learning

A *deep learning approach* had a higher correlation with GPA than other non-cognitive study factors (see Table 7) and was the only non-cognitive factor of learning with a statistically significant difference in the mean score across all three GPA bands of high-risk, medium-risk, and low-risk participants (Appendix, Table 13). However, only half of classification models used one of the three factors of learning approach: BPNN, SVM, LR, and a voting ensemble. Volet (1996) found that goal setting influences self-regulation, which in turn influences learning approach adopted. Similarly in this study, a *deep learning approach* had the strongest correlations with *intrinsic goal orientation* ( $r=0.417$ ), *metacognitive self-regulation* ( $r=0.431$ ), and *study effort* ( $r=0.360$ ) (see Appendix, Table 15). With the exception of a bagging ensemble, all models that found approaches to learning to be unimportant showed that either *goal orientation* or *study effort* were significant. Additionally, apart from a voting ensemble, no model used both learning approaches and *study effort*, although a number of models used learning approaches with learning goals.

#### 7.3.7 Learner modality

While awareness of learner modality by student and lecturer can improve the student-learning experience (Duffin & Gray, 2009a; Gilakjani, 2012), there is no evidence to suggest that learner modality is predictive of academic performance (Gilakjani, 2012; Kablan, 2016). Correlation results (see Table 7)

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

and analysis of group differences (see Table 13) concurred with this observation for both *visual* and *auditory* modalities. However, the *kinaesthetic modality* (learn by doing) had a weak but statistically significant correlation with GPA ( $r=-0.059$ ,  $p<0.05$ ). In addition, there were statistically significant group differences for *kinaesthetic modality* by GPA band ( $p<0.05$ ). *Kinaesthetic modality* was significant in six of the eight classification models. Kinaesthetic learners were more likely to be male, more likely to have registered poor prior academic performance, and displayed a weak but negative correlation with factors of effective learning disposition. Further work is needed to determine the real importance of *kinaesthetic modality*, i.e., if it was used as a proxy for poor learning disposition, or was itself indicative of failure.

## 7.4 Malleable Learner Dispositions

Comparison of models with and without non-cognitive factors of learning (personality, motivation, self-regulation, approaches to learning, and learner modality) suggested that the addition of non-cognitive factors of learning provided limited improvement in predictive accuracy in spite of their significant correlations with GPA. Therefore, their value in learner profiling at student induction merits consideration. Several studies have reported learning disposition is malleable. For example, Miller-Reilly (2006) evidenced that teaching approaches had changed adult learners' self-efficacy in mathematics. Similarly, a meta-analysis of studies on self-regulation reported improvements in self-regulation, and consequently learning goals, following self-regulation training and support (Winters et al., 2008). It could be argued that all students at risk of failing require further support in developing effective learning dispositions regardless of individual profiles. However, a profiling tool facilitates feedback to both students and lecturers that may support other interventions. The profiler used in this study gave immediate feedback to students on their learner profile. Duffin and Gray (2009b) found that 56% of students understood their learning profile based on online feedback, and this rose to 83% when profiling was followed up by explanatory workshops. Arnold and Pistilli (2012) reported a 6.4 percentage point decrease in grades D, F, and withdrawals amongst users of their Course Signals tool that provided both early warnings and suggested improvement strategies to at-risk students starting in week two. However, Jayaprakash et al. (2014) found that simply making students aware that they may be at risk of failing significantly increased the numbers passing and the number of withdrawals, but providing further supports did not effect additional changes in either measure. Therefore, further work is needed to assess the impact of timely feedback on learner disposition, specifically on subsequent optimal use of that feedback.

## 8 CONCLUSION

Empirical models of learning developed in this study predicted first-year students at risk of failing with an accuracy of 72% when applied to a new student cohort. The dataset was diverse in terms of age, academic discipline, and assessment strategies used ( $n=1,207$ ). Informed by a review of factors predictive of academic performance in tertiary education, study factors related to prior academic performance, personality, motivation, learning strategies, learner modality, age, and gender. The



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

twenty-four study factors used were measured prior to or during first-year student enrollment. Model accuracies were comparable across the eight classification algorithms used, namely  $k$ -nearest neighbour ( $k$ -NN), a voting ensemble, support vector machine (SVM),  $k$ -NN bagging ensemble, decision tree (DT), back-propagation neural network (BPNN), naïve Bayes (NB), and logistic regression (LR). As expected, 10-fold cross validation model accuracies were higher than models applied to a different student cohort; however, the increase in model accuracy was statistically significant for SVM only. Subgroups smaller than  $n=900$  under-represented patterns in the dataset.

Analysis of misclassifications showed that *fails* misclassified as *pass* did not differ in learning disposition or prior academic performance from those correctly classified as *pass*. Factors measurable later in the semester, such as academic and social integration, economic pressures, and teaching methods, may explain misclassifications. Further work is needed to determine potential improvements in model accuracy in an Institute of Technology student cohort if data gathered after student induction were included in the model.

Attributes that were significant in accurate prediction of students at risk of failing across a range of courses included:

- Age: The study sample had an age range of [18,60]. Younger students had a greater risk of failing in their first year of study.
- Aggregates of prior academic performance: In particular, an aggregate of mathematics, science, and business related subjects was found to be a stronger predictor of year 1 students at risk of failing compared to other prior academic performance aggregates.
- Factors of motivation, particularly self-efficacy and intrinsic goal orientation.
- Learning approach: A shallow or a strategic learning approach was indicative of students at risk of failing.
- Openness, indicating a creative, inquisitive temperament, was indicative of a passing grade.
- A kinaesthetic modality (preference for learn by doing) was indicative of students at risk of failing.

Self-regulation factors were not found to be significant once learning goals and approaches to learning were considered. Similarly, conscientiousness did not improve model accuracy over and above other factors of learning. Respective correlations between study factors and GPA were not indicative of factors significant in classification models of students at risk of failing. For example, both openness and kinaesthetic learner modality were significant in a number of classification models but had poor correlations with GPA. Conclusions from this study that openness and kinaesthetic learner modality were significant predictors of academic performance were not widely observed in other studies, particularly the significance of learner modality. Therefore, further work is needed to determine if their importance in models of learning, as reported in this study, generalizes to other student cohorts.

The primary value of non-cognitive factors of learning in this study was to distinguish the learning profile of students at risk of failing from the learning profile of students that passed, rather than provide



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

improvement in model predictive accuracy. It has been argued that non-cognitive factors of motivation, self-regulation, and approaches to learning are malleable, and key to an effective learning disposition, which in turn should be a valued learning outcome of courses in tertiary education. Further work is needed to evaluate subsequent benefits of learner profiling during student induction, both for the student, and for first-year mentoring and support programmes.

## ACKNOWLEDGEMENTS

The authors would like to thank the Institute of Technology Blanchardstown ([www.itb.ie](http://www.itb.ie)) for its support in facilitating this research, and the staff at the National Learning Network Assessment Services ([www.nln.ie](http://www.nln.ie)) for their assistance in administering questionnaires during student induction.

## REFERENCES

- ACT. (2012). *2012 Retention Completion Summary Tables* (Technical report). Retrieved from [www.act.org](http://www.act.org)
- Ali, L., Hatala, M., Gašević, D., & Winne, P. H. (2014). Leveraging MSLQ data for predicting student achievement goal orientations. *Journal of Learning Analytics*, 1(3), 157–160. Retrieved from <https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4215>
- Allick, J., & Realo, A. (1997). Intelligence, academic abilities, and personality. *Personality and Individual Differences*, 23(5), 809–814 [http://dx.doi.org/10.1016/S0191-8869\(97\)00103-7](http://dx.doi.org/10.1016/S0191-8869(97)00103-7)
- Apter, M. J. (1989). *Reversal theory: Motivation, emotion and personality*. London: Routledge.
- Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. *Proceedings of the 2<sup>nd</sup> International Conference on Learning Analytics and Knowledge* (LAK '12), 267–270. <http://dx.doi.org/10.1145/2330601.2330666>
- Baker, R. S. J. d., & Yacef, K. (2010). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3–17. Retrieved from <http://educationaldatamining.org/JEDM/index.php/JEDM/article/view/8>
- Baumann, K. (2003). Cross-validation as the objective function for variable-selection techniques. *Trends in Analytical Chemistry*, 22(6), 395–406. [http://dx.doi.org/10.1016/S0165-9936\(03\)00607-1](http://dx.doi.org/10.1016/S0165-9936(03)00607-1)
- Bergin, S. (2006). *Statistical and machine learning models to predict programming performance* (Unpublished doctoral dissertation). National University of Ireland, Maynooth.
- Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, 17(1), 69–81. <http://dx.doi.org/10.1016/j.lindif.2007.02.001>
- Biggs, J., Kember, D., & Leung, D. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71, 133–149. <http://dx.doi.org/10.1007/BF03173004>
- Bruinsma, M. (2004). Motivation, cognitive processing and achievement in higher education. *Learning and Instruction*, 14(6), 549–568. <http://dx.doi.org/10.1016/j.learninstruc.2004.09.001>
- Buckingham Shum, S., & Deakin Crick, R. (2012). Learning dispositions and transferable competencies: Pedagogy, modelling and learning analytics. *Proceedings of the 2<sup>nd</sup> International Conference on*

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

- Learning Analytics and Knowledge* (LAK '12), 92–101. <http://dx.doi.org/10.1145/2330601.2330629>
- Burisch, M. (1997). Test length and validity revisited. *European Journal of Personality*, 11(4), 303–315.
- Campbell, J., deBlois, P., & Oblinger, D. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review*, 42(4), 40–57.
- Carpenter, J., & Bithell, J. (2000). Bootstrap confidence intervals: When, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, 19(9), 1141–1164.
- Cassidy, S. (2011). Exploring individual differences as determining factors in student academic achievement in higher education. *Studies in Higher Education*, 37(7), 1–18. <http://dx.doi.org/10.1080/03075079.2010.545948>
- Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences*, 44(7), 1596–1603. <http://dx.doi.org/10.1016/j.paid.2008.01.003>
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, Special Issue on State of the Art in TEL*, 318–331.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kelelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16(1), 321–357.
- Colby, J. (2005). Attendance and attainment: A comparative study. *Innovation in Teaching and Learning in Information and Computer Sciences*, 4(2), 1–13.
- Conrad, M. A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, 40(3), 339–346. <http://dx.doi.org/10.1016/j.jrp.2004.10.003>
- Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology*, 51, 171–200. <http://dx.doi.org/10.1146/annurev.psych.51.1.171>
- De Clercq, M., Galand, B., & Frenay, M. (2013). Chicken or the egg: Longitudinal analysis of the causal dilemma between goal orientation, self-regulation and cognitive processing strategies in higher education. *Studies in Educational Evaluation*, 39(1), 4–13. <http://dx.doi.org/10.1016/j.stueduc.2012.10.003>
- de Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality*, 10(5), 303–336.
- Deakin Crick, R., & Goldspink, C. (2014). Learning dispositions, self-theories and student engagement. *British Journal of Educational Studies*, 62(1), 19–35. <http://dx.doi.org/10.1080/00071005.2014.904038>
- Dekker, G., Pechenizkiy, M., & Vleeshouwers, J. (2009). Predicting students drop out: A case study. In T. Barnes, M. Desmarais, C. Romero, & S. Ventura (Eds.), *Proceedings of the 2<sup>nd</sup> Annual Conference on Educational Data Mining (EDM '09)*, 41–50. Retrieved from <http://www.educationaldatamining.org/conferences/index.php/EDM/2009/paper/download/1467/1433>
- Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

- algorithms. *Journal of Neural Computation*, 10(7), 1895–1923. <http://dx.doi.org/10.1162/089976698300017197>
- Diseth, Á. (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learning and Individual Differences*, 21(2), 191–195. <http://dx.doi.org/10.1016/j.lindif.2011.01.003>
- Dollinger, S. J., Matyja, A. M., & Huber, J. L. (2008). Which factors best account for academic success: Those which college students can control or those they cannot? *Journal of Research in Personality*, 42(4), 872–885. <http://dx.doi.org/10.1016/j.jrp.2007.11.007>
- Drachler, H., & Greller, W. (2012). The pulse of learning analytics: Understandings and expectations from the stakeholders. *Proceedings of the 2<sup>nd</sup> International Conference on Learning Analytics and Knowledge (LAK '12)*, 120–129. <http://dx.doi.org/10.1145/2330601.2330634>
- Duff, A., Boyle, E., Dunleavy, K., & Ferguson, J. (2004). The relationship between personality, approach to learning and academic performance. *Personality and Individual Differences*, 36(2), 1907–1920. <http://dx.doi.org/10.1016/j.paid.2003.08.020>
- Duffin, D., & Gray, G. (2009a). Accommodating learner diversity in the classroom, *Proceedings of the 3<sup>rd</sup> International Technology, Education and Development Conference (INTED 2009)*, 9–11 March 2009, Valencia, Spain (pp. 4629–4635). Retrieved from <https://library.iated.org/view/DUFFIN2009ACC>
- Duffin, D., & Gray, G. (2009b). Using ICT to enable inclusive teaching practices in higher education. In P. L. Emiliani, L. Burzagli, A. Como, F. Gabbanini, & A.-L. Salminen (Eds.), *Assistive Technology Research Series*, 25, 640–645.
- Elliot, A., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, 70(3), 461–475.
- Entwistle, N. (2005). Contrasting perspectives in learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education*, 3<sup>rd</sup> (Internet) edition (pp. 3–22). Edinburgh: University of Edinburgh, Centre for Teaching, Learning and Assessment. Retrieved from <http://www.tla.ed.ac.uk/resources/EoL.html>
- Eppler, M. A., & Harju, B. L. (1997). Achievement motivation goals in relation to academic performance in traditional and nontraditional college students. *Research in Higher Education*, 38(5), 557–573. <http://dx.doi.org/10.1023/A:1024944429347>
- Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: The roles of personality, intelligence, and application. *Personality and Individual Differences*, 34(7), 1225–1243. [http://dx.doi.org/10.1016/S0191-8869\(02\)00111-3](http://dx.doi.org/10.1016/S0191-8869(02)00111-3)
- Fleming, N. D. (1995). I'm different, not dumb: Modes of presentation (VARK) in the tertiary classroom. In A. Zelmer (Ed.), *Research and Development in Higher Education, Proceedings of the 1995 Annual Conference of the Higher Education and Research Development Society of Australasia (HERDSA 1995)*, 4–8 July 1995, Rockhampton, Australia (Vol. 18, pp. 308–313).
- Galesic, M., & Bosnjak, M. (2009). Effect of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, 73(2), 349–360. <http://dx.doi.org/10.1093/poq/nfp031>

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

- Ganyaupfu, E. M. (2013). Teaching methods and students' academic performance. *International Journal of Humanities and Social Science Invention*, 2(9), 29–35.
- Gilakjani, A. P. (2012). A match or mismatch between learning styles of the learners and teaching styles of the teachers. *International Journal of Modern Education and Computer Science*, 11, 51–60.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40, 84–96. <http://dx.doi.org/10.1016/j.jrp.2005.08.007>
- Gray, G., McGuinness, C., & Owende, P. (2013). An investigation of psychometric measures for modelling academic performance in tertiary education. In S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.), *Proceedings of the 6<sup>th</sup> International Conference on Educational Data Mining (EDM 2013)*, 6–9 July, Memphis, TN, USA (pp. 240–243). Springer.
- Gray, G., McGuinness, C., & Owende, P. (2014). Non-cognitive factors of learning as predictors of academic performance in tertiary education. In S. Ritter & S. Fancsali (Eds.), *Workshop on Non-Cognitive Factors and Personalization for Adaptive Learning (NCFPAL 2014)* at the 7<sup>th</sup> International Conference on Educational Data Mining (EDM 2014), CEUR Workshop Proceedings, (pp. 107–114): International Educational Data Mining Society.
- Gray, G., McGuinness, C., Owende, P., & Carthy, A. (2014). A review of psychometric data analysis and applications in modelling of academic achievement in tertiary education. *Journal of Learning Analytics*, 1(1), 75–106.
- Hair, J. F. J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective, 7th ed.* New York: Pearson Education.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Association of Physics Teachers*, 66, 64–74. <http://dx.doi.org/10.1119/1.18809>
- Hall, M. A., & Homes, G. (2003). Benchmarking attributes selection techniques for discrete class data mining. *IEEE Transactions on Knowledge and Data Engineering*, 15(6), 1437–1447. <http://dx.doi.org/10.1109/TKDE.2003.1245283>
- Herzog, S. (2006). Estimating student retention and degree-completion time: Decision trees and neural networks vis-à-vis regression. *New Directions for Institutional Research*, 131, 17–33. <http://dx.doi.org/10.1002/ir.185>
- Hoskins, S. L., Newstead, S. E., & Dennis, I. (1997). Degree performance as a function of age, gender, prior qualifications and discipline studied. *Assessment and Evaluation in Higher Education*, 22(3), 317–328. <http://dx.doi.org/10.1080/0260293970220305>
- Jayaprakash, S. M., Moody, E. W., Lauria, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), 6–47. Retrieved from <https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3249>
- Kablan, Z. (2016). The effect of manipulatives on mathematics achievement across different learning styles. *Educational Psychology*, 36(2), 277–296.

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

<http://dx.doi.org/10.1080/01443410.2014.946889>

Kang, Y., & Harring, J. R. (2015). Reexamining the impact of non-normality in two-group comparison procedures. *Journal of Experimental Education*, 83(2), 147–174.

<http://dx.doi.org/10.1080/00220973.2013.876605>

Kappe, R., & van der Flier, H. (2010). Using multiple and specific criteria to assess the predictive validity of the big five personality factors on academic performance. *Journal of Research in Personality*, 44(1), 142–145. <http://dx.doi.org/10.1016/j.jrp.2009.11.002>

Kaufman, J. C., Agars, M. D., & Lopez-Wagner, M. C. (2008). The role of personality and motivation in predicting early college academic success in non-traditional students at a Hispanic-serving institution. *Learning and Individual Differences*, 18(4), 492–496.

<http://dx.doi.org/10.1016/j.lindif.2007.11.004>

Knight, S., Buckingham Shum, S., & Littleton, K. (2013). Epistemology, pedagogy, assessment and learning analytics. *Proceedings of the 3<sup>rd</sup> International Conference on Learning Analytics and Knowledge* (LAK '13), 75–84. Retrieved from

<https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3538>

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The big five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, 51(4), 472–477.

<http://dx.doi.org/10.1016/j.paid.2011.04.019>

Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences*, 25, 67–72.

<http://dx.doi.org/10.1016/j.lindif.2013.01.005>

Kundel, H. L., & Polansky, M. (2003). Measurement of observer agreement. *Radiology*, 228(2), 303–308.

<http://dx.doi.org/10.1148/radiol.2282011860>

Lauria, E. J. M., Moody, E. W., Jayaprakash, S. M., Jonnalagadda, N., & Baron, J. D. (2013). Open academic analytics initiative: Initial research findings. *Proceedings of the 3<sup>rd</sup> International Conference on Learning Analytics and Knowledge* (LAK '13), 150–154.

<http://dx.doi.org/10.1145/2460296.2460325>

Marton, F., & Säljö, R. (2005). Approaches to learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education*, 3<sup>rd</sup> (Internet) edition (pp. 36–58). Edinburgh: University of Edinburgh, Centre for Teaching, Learning and Assessment. Retrieved from <http://www.tla.ed.ac.uk/resources/EoL.html>

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105(1), 156–166. <http://dx.doi.org/10.1037/0033-2909.105.1.156>

Miller-Reilly, B. (2006). *Affective change in adult students in second chance mathematics courses: Three different teaching approaches* (Unpublished doctoral dissertation). University of Auckland, New Zealand.

Milne, J., Jeffrey, L. M., Suddaby, G., & Higgins, A. (2012). Early identification of students at risk of failing. *Proceedings of the 29<sup>th</sup> Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (ASCILITE 2012), (Vol. 1, pp. 25–28).

Minaei-Bidgoli, B., Kashy, D. A., Kortemeyer, G., & Punch, W. F. (2003). Predicting student performance:



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

- An application of data mining methods with educational web-based system lon-capa. *Proceedings of the 33<sup>rd</sup> ASEE/IEEE Frontiers in Education Conference*, 13–18.
- Mirriahi, N., Gašević, D., Long, P., & Dawson, S. (2014). Scientometrics as an important tool for the growth of the field of learning analytics. *Journal of Learning Analytics*, 1(2), 1–4. Retrieved from <https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4116>
- Mitchell, T. M. (2015). Generative and discriminative classifiers: Naïve Bayes and logistic regression. In *Machine Learning*, 2<sup>nd</sup> ed., McGraw Hill. Retrieved from <http://www.cs.cmu.edu/~tom/mlbook/NBayesLogReg.pdf>
- Mooney, O., Patterson, V., O'Connor, M., & Chantler, A. (2010). *A Study of Progression in Higher Education: A Report by the Higher Education Authority* (Technical report). Higher Education Authority, Ireland. Retrieved from [http://www.heai.ie/sites/default/files/study\\_of\\_progression\\_in\\_irish\\_higher\\_education\\_2010.pdf](http://www.heai.ie/sites/default/files/study_of_progression_in_irish_higher_education_2010.pdf)
- Moran, M. A., & Crowley, M. J. (1979). The leaving certificate and first year university performance. *Journal of Statistical and Social Enquiry in Ireland*, 24(1), 231–266.
- Naderi, H., Abdullah, H. T., Sharir, J., & Kumar, V. (2009). Creativity, age and gender as predictors of academic achievement among undergraduate students. *Journal of American Science*, 5(5), 101–112.
- Ng, A. Y., & Jordan, M. I. (2001). On discriminative vs. generative classifiers: A comparison of logistic regression and naïve Bayes. *Proceedings of the 15<sup>th</sup> Conference on Advances in Neural Information Processing Systems (NIPS 2001)*, pp. 841–848.
- OECD. (2013). *Education At a Glance 2013: OECD Indicators*. OECD Publishing. Retrieved from <http://dx.doi.org/10.1787/eag-2013-en>
- Patterson, V., Carroll, D., & Harvey, V. (2014). *HEA Key Facts and Figures: Higher Education 2012/13* (Technical report). Higher Education Authority, Ireland. Retrieved from [http://www.heai.ie/sites/default/files/key\\_facts\\_and\\_figures\\_2012-13.pdf](http://www.heai.ie/sites/default/files/key_facts_and_figures_2012-13.pdf)
- Pérez-Martínez, J. E., García-García, M. J., Perdomo, W. H., & Villamide-Díaz, M. J. (2009). Analysis of the results of the continuous assessment in the adaptation of the Universidad Politécnica de Madrid to the European Higher Education Area. In R. Hadgraft & L. Mann (Eds.), *Proceedings of the Research in Engineering Education Symposium (REES2009)*. Retrieved from [http://rees2009.pbworks.com/f/rees2009\\_submission\\_46.pdf](http://rees2009.pbworks.com/f/rees2009_submission_46.pdf)
- Pintrich, P., Smith, D., Garcia, T., & McKeachie, W. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire*. [Technical Report 91-B-004]. The Regents of the University of Michigan. Retrieved from <http://files.eric.ed.gov/fulltext/ED338122.pdf>
- Provost, F., Jensen, D., & Oates, T. (1999). Efficient progressive sampling. *Proceedings of the 5<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '99)*, 23–31. <http://dx.doi.org/10.1145/312129.312188>
- Rice, J. A. (1995). *Mathematical statistics and data analysis*, 2nd ed. Belmont, CA: Duxbury Press.
- Robbins, S. B., Lauver, K., Le, H., Davis, D., & Langley, R. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130(2), 261–288.



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

<http://dx.doi.org/10.1037/0033-2909.130.2.261>

- Romero, C., Ventura, S., Espejo, P. G., & Hervás, C. (2008). Data mining algorithms to classify students. In R. S. J. d. Baker, T. Barnes, & J. E. Beck (Eds.), *Proceedings of the 1<sup>st</sup> International Conference on Educational Data Mining (EDM 2008)*, 8–17. Retrieved from <http://www.educationaldatamining.org/conferences/index.php/EDM/2008/paper/download/756/722>
- Roll, I., & Winne, P. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, 2(1), pp. 7–12. Retrieved from <https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4491>
- Sachin, B. R., & Vijay, S. M. (2012). A survey and future vision of data mining in educational field. *Proceedings of the 2<sup>nd</sup> International Conference on Advanced Computing Communication Technologies (ACCT12)*, (pp. 96–100). <http://dx.doi.org/10.1109/ACCT.2012.14>
- Shute, V., & Ventura, M. (2013). *Stealth assessment: Measuring and supporting learning in video games*. Cambridge, Massachusetts: MIT Press.
- Siemens, G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice. *Proceedings of the 2<sup>nd</sup> International Conference on Learning Analytics and Knowledge (LAK '12)*, 4–8. <http://dx.doi.org/10.1145/2330601.2330605>
- Siemens, G., & Baker, R. S. J. d. (2012). Learning analytics and educational data mining: Towards communication and collaboration. *Proceedings of the 2<sup>nd</sup> International Conference on Learning Analytics and Knowledge (LAK '12)*, 252–254. <http://dx.doi.org/10.1145/2330601.2330661>
- Sins, P. H. M., van Joolingen, W. R., Savelsbergh, E. R., & van Hout-Wolters, B. (2008). Motivation and performance within a collaborative computer-based modeling task: Relations between students' achievement goal orientation, self-efficacy, cognitive processing, and achievement. *Contemporary Educational Psychology*, 33(1), 58–77. <http://dx.doi.org/10.1016/j.cedpsych.2006.12.004>
- Smith, Z. R., & Wells, C. S. (2006). Central limit theorem and sample size. Paper presented at the annual meeting of the Northeastern Educational Research Association, 18–20 October 2006, Kerhonkson, New York. Retrieved from [http://www.umass.edu/rempp/Papers/Smith&Wells\\_NERA06.pdf](http://www.umass.edu/rempp/Papers/Smith&Wells_NERA06.pdf)
- Steel, P., & Konig, C. J. (2006). Integrating theories of motivation. *Academy of Management Review*, 31(4), 889–913. <http://dx.doi.org/10.5465/AMR.2006.22527462>
- Superby, J. F., Vandamme, J. P., & Meskens, N. (2006). Determination of factors influencing the achievement of the first-year university students using data mining methods. *Proceedings of the Workshop on Educational Data Mining at the 8<sup>th</sup> International Conference on Intelligent Tutoring Systems (ITS 2006)*, 37–44. Retrieved from <http://www.educationaldatamining.org/IEDMS/sites/default/files/files/EDMITS2006/superby.pdf>
- Swanberg, A. B., & Martinsen, Ø. L. (2010). Personality, approaches to learning and achievement. *Educational Psychology*, 30(1), 75–88. <http://dx.doi.org/10.1080/01443410903410474>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's Alpha. *International Journal of Medical*

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

- Education*, 2, 53–55. <http://dx.doi.org/10.5116/ijme.4dfb.8dfd>
- Tempelaar, D. T., Cuypers, H., van de Vrie, E., Heck, A., & van der Kooij, H. (2013). Formative assessment and learning analytics. *Proceedings of the 3<sup>rd</sup> International Conference on Learning Analytics and Knowledge* (LAK '13), 205–209. <http://dx.doi.org/10.1145/2460296.2460337>
- Thai-Nghe, N. Janecek, P., & Haddawy, P. (2007). A comparative analysis of techniques for predicting academic performance. *Proceedings of the 37<sup>th</sup> ASEE/IEEE Frontiers in Education Conference*, 10–13 October 2007, Milwaukee, WI, USA (pp. T2G-7–T2G-12).
- Tinto, V. (2006). Research and practice of student retention: What next? *Journal of College Student Retention*, 8(1), 1–19. <http://dx.doi.org/10.2190/4YNU-4TMB-22DJ-AN4W>
- Tishman, S., Jay, E., & Perkins, D. N. (1993). Teaching thinking disposition: From transmission to enculturation. *Theory into Practice*, 32(3), 147–153. <http://dx.doi.org/10.1080/00405849309543590>
- Volet, S. E. (1996). Cognitive and affective variables in academic learning: The significance of direction and effort in students' goals. *Learning and Instruction*, 7(3), 235–254. [http://dx.doi.org/10.1016/S0959-4752\(96\)00027-8](http://dx.doi.org/10.1016/S0959-4752(96)00027-8)
- Weiss, G. M. (2004). Mining with rarity: A unifying framework. *SIGKDD Explorations*, 6(1), 7–19.
- Wigfield, A., Eccles, J. S., & Pintrich, P. (1996). *Handbook of educational psychology* (pp. 148–185). New York: Simon & Schuster/Macmillan.
- Winters, F. I., Greene, J. A., & Costich, C. M. (2008). Self-regulated learning within computer-based learning environments a critical analysis. *Educational Psychology Review*, 20(4), 429–444. <http://dx.doi.org/10.1007/s10648-008-9080-9>
- Wise, A. F., & Shaffer, D. W. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics*, 2(2), 5–13. <http://dx.doi.org/10.18608/jla.2015.22.2>
- Yang, J., & Honavar, V. (1998). Feature subset selection using a genetic algorithm. *IEEE Intelligent Systems*, 13(2), 44–49.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 3–17. [http://dx.doi.org/10.1207/s15326985ep2501\\_2](http://dx.doi.org/10.1207/s15326985ep2501_2)

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372.  
<http://dx.doi.org/10.18608/jla.2016.32.20>

## APPENDIX

**Table 12: Leaving certificate grades and their corresponding CAO points.**

Mark(%)	Grade	Higher level	Ordinary level	Foundation level
90–100	A1	100	60	20
85–89.99	A2	90	50	15
80–84.99	B1	85	45	10
75–79.99	B2	80	40	5
70–74.99	B3	75	35	0
65–69.99	C1	70	30	0
60–64.99	C2	65	25	0
55–59.99	C3	60	20	0
50–54.99	D1	55	15	0
45–49.99	D2	50	10	0
40–44.99	D3	45	5	0
<40	E, F, NG	0	0	0

(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372.  
<http://dx.doi.org/10.18608/jla.2016.32.20>

**Table 13: Group differences by GPA band.**

Study Factor	Range	2.5<GPA≤4 n=558	2.0≤GPA<2.5 n=190	0≤GPA<2 n=459	p
CAO points	[0,600]	223.5 ± 137.1	221.1 ± 117.6	212.2 ± 91.4	***
Leaving cert average	[0,100]	33.4 ± 20.6	33.1 ± 17.4	30.9 ± 13.5	***
Mathematics	[0,100]	21.6 ± 17.1	21.4 ± 15.4	17.1 ± 13.1	***
English	[0,100]	39.1 ± 26.3	40.6 ± 24.0	38.3 ± 21.0	***
Applied average	[0,100]	24.8 ± 29.4	26.4 ± 28.6	27.4 ± 26.3	
Methodical average	[0,100]	29.1 ± 20.8	28.8 ± 18.2	23.8 ± 14.7	***
Humanities average	[0,100]	34.2 ± 22.3	34.0 ± 19.0	33.0 ± 15.7	***
Conscientiousness	[0,10]	6.2 ± 1.5	5.8 ± 1.4	5.7 ± 1.6	***
Openness	[0,10]	6.2 ± 1.3	5.9 ± 1.2	6.0 ± 1.4	**
Self-efficacy	[0,10]	7.0 ± 1.4	6.9 ± 1.3	6.7 ± 1.5	***
Extrinsic goals	[0,10]	8.0 ± 1.3	7.8 ± 1.3	7.6 ± 1.4	***
Intrinsic goals	[0,10]	7.3 ± 1.3	6.9 ± 1.4	6.9 ± 1.4	***
MC self-regulation	[0,10]	6.0 ± 1.3	5.9 ± 1.3	5.7 ± 1.4	**
Study effort	[0,10]	6.2 ± 1.7	5.9 ± 1.7	5.6 ± 1.8	***
Study time	[0,10]	6.3 ± 2.2	6.4 ± 2.3	5.9 ± 2.4	*
Deep learner	[0,10]	6.0 ± 2.9	5.3 ± 2.8	4.7 ± 2.8	***
Shallow learner	[0,10]	1.0 ± 1.8	1.4 ± 1.9	1.7 ± 2.1	***
Strategic learner	[0,10]	3.1 ± 2.4	3.5 ± 2.5	3.8 ± 2.5	***
Gender	[0,10]	0.4 ± 0.5	0.5 ± 0.5	0.3 ± 0.5	***
Group work	[0,10]	6.2 ± 3.5	7.2 ± 2.9	6.8 ± 3.3	***
Age	[0,10]	25.1 ± 8.4	22.5 ± 6.7	21.4 ± 5.5	***
Visual	[0,10]	7.3 ± 2.0	7.3 ± 2.0	7.0 ± 2.1	
Auditory	[0,10]	3.2 ± 2.2	3.0 ± 2.0	3.1 ± 2.2	
Kinaesthetic	[0,10]	4.5 ± 2.4	4.7 ± 2.3	4.9 ± 2.4	*

MC=Metacognitive; \*p<0.05; \*\*p<0.01; \*\*\*p<0.001; Colours differentiate heat map scales.

**Table 14: Group differences by age and gender.**

Study Factor	Range	18–23, n=875	24–28, n=131	29+, n=21	p	Male, n=713	Female, n=494	p
CAO points	[0,600]	266.0 ± 76.0	147.9 ± 127.6	59.9 ± 102.2	***	209.9 ± 117.4	231.7 ± 119.1	***
Average leaving cert	[0,100]	39.1 ± 11.8	22.3 ± 19.2	9.7 ± 16.1	***	31.0 ± 17.5	34.4 ± 17.8	***
Mathematics	[0,100]	23.4 ± 13.7	15.4 ± 17.3	7.4 ± 14.9		19.1 ± 15.5	21.1 ± 15.6	*
English	[0,100]	47.6 ± 18.2	27.5 ± 25.4	9.3 ± 17.6	***	36.8 ± 24	42.3 ± 23.6	***
Applied average	[0,100]	31.5 ± 28.2	18.6 ± 26.5	7.0 ± 17.2	**	27.1 ± 28.5	24.5 ± 27.4	*
Methodical average	[0,100]	32.5 ± 15.5	19.0 ± 18.8	8.3 ± 15.5	***	24.6 ± 17.6	30.5 ± 18.9	***
Humanities average	[0,100]	41.1 ± 13.9	22.6 ± 20.1	8.8 ± 15.8	***	32.2 ± 19.3	35.9 ± 19.6	***
Conscientiousness	[0,10]	5.8 ± 1.5	6.0 ± 1.5	6.5 ± 1.4	***	6.0 ± 1.5	5.9 ± 1.6	
Openness	[0,10]	6.0 ± 1.3	6.5 ± 1.3	6.1 ± 1.3	***	6.0 ± 1.3	6.1 ± 1.2	
Self-efficacy	[0,10]	6.8 ± 1.4	7.0 ± 1.4	7.0 ± 1.5		6.9 ± 1.4	6.8 ± 1.4	
Extrinsic goals	[0,10]	7.7 ± 1.4	7.9 ± 1.3	8.1 ± 1.4	**	7.8 ± 1.4	7.9 ± 1.3	
Intrinsic goals	[0,10]	6.9 ± 1.3	7.5 ± 1.2	7.8 ± 1.3	***	7.1 ± 1.4	7.1 ± 1.4	
Metacognitive self-regulation	[0,10]	5.7 ± 1.3	6.1 ± 1.3	6.6 ± 1.2	***	5.9 ± 1.4	5.9 ± 1.4	
Study effort	[0,10]	5.7 ± 1.8	6.2 ± 1.7	6.8 ± 1.5	***	5.9 ± 1.8	6.0 ± 1.7	
Study time	[0,10]	6.2 ± 2.3	5.8 ± 2.4	6.5 ± 2.2	*	6.0 ± 2.3	6.4 ± 2.3	**
Deep learner	[0,10]	4.8 ± 2.8	6.5 ± 2.8	7.1 ± 2.5	***	5.2 ± 3	5.7 ± 2.8	**
Shallow learner	[0,10]	1.6 ± 2.1	0.8 ± 1.5	0.6 ± 1.3	***	1.5 ± 2.1	1.0 ± 1.7	***
Strategic learner	[0,10]	3.7 ± 2.5	2.8 ± 2.3	2.4 ± 2.1	***	3.4 ± 2.5	3.4 ± 2.5	
Gender	[0,10]	0.4 ± 0.5	0.4 ± 0.5	0.4 ± 0.5				
Group work	[0,10]	6.6 ± 3.3	6.1 ± 3.6	6.5 ± 3.4		6.5 ± 3.3	6.7 ± 3.4	
Age	[0,10]	19.6 ± 1.2	25.7 ± 1.3	37.6 ± 7.0	***	23.5 ± 7.7	22.9 ± 6.8	
Visual	[0,10]	7.1 ± 2.1	7.3 ± 2.0	7.3 ± 2.1		7.2 ± 2.1	7.1 ± 2.0	
Auditory	[0,10]	3.1 ± 2.1	3.1 ± 2.2	3.2 ± 2.2		2.8 ± 2.2	3.7 ± 2.1	***
Kinaesthetic	[0,10]	4.7 ± 2.4	4.6 ± 2.5	4.4 ± 2.4		5.0 ± 2.5	4.2 ± 2.2	***
GPA	[0,4]	1.9 ± 1.1	2.3 ± 1.0	2.7 ± 0.9	***	2.0 ± 1.1	2.2 ± 1.0	***

\*p&lt;0.05; \*\*p&lt;0.01; \*\*\*p&lt;0.001; Colours differentiate heat map scales.

**Table 15: Correlations for non-cognitive factors of learning.**

	GPA	Conscientiousness	Openness	Self-efficacy	Extrinsic goal	Intrinsic goal
Conscientiousness	0.150 [0.09, 0.21]					
Openness	0.084 [0.03, 0.14]	0.032 [-0.03, 0.09]				
Self-efficacy	0.120 [0.06, 0.18]	0.313 [0.26, 0.36]	0.178 [0.12, 0.23]			
Extrinsic goal	0.124 [0.07, 0.18]	0.280 [0.22, 0.33]	0.049 [-0.01, 0.11]	0.308 [0.25, 0.36]		
Intrinsic goal	0.149 [0.09, 0.21]	0.334 [0.28, 0.39]	0.316 [0.26, 0.37]	0.421 [0.37, 0.47]	0.381 [0.33, 0.43]	
MC self-regulation	0.130 [0.08, 0.18]	0.515 [0.47, 0.56]	0.101 [0.04, 0.16]	0.409 [0.36, 0.46]	0.298 [0.24, 0.35]	0.429 [0.38, 0.48]
Study effort	0.187 [0.14, 0.24]	0.450 [0.39, 0.50]	0.064 [0.01, 0.13]	0.334 [0.28, 0.39]	0.232 [0.17, 0.29]	0.330 [0.27, 0.38]
Study time	0.101 [0.04, 0.16]	0.396 [0.35, 0.44]	0.009 [-0.05, 0.06]	0.259 [0.21, 0.31]	0.175 [0.12, 0.23]	0.227 [0.17, 0.28]
Deep learner	0.234 [0.18, 0.29]	0.352 [0.30, 0.40]	0.209 [0.15, 0.26]	0.273 [0.22, 0.32]	0.158 [0.10, 0.21]	0.417 [0.37, 0.47]
Strategic learner	-0.158 [-0.22, -0.1]	-0.167 [-0.22, -0.11]	-0.174 [-0.23, -0.12]	-0.158 [-0.21, -0.1]	-0.012 [-0.06, 0.04]	-0.274 [-0.33, -0.22]
Shallow learner	-0.146 [-0.21, -0.09]	-0.330 [-0.38, -0.28]	-0.096 [-0.15, -0.04]	-0.221 [-0.28, -0.16]	-0.234 [-0.29, -0.17]	-0.294 [-0.36, -0.24]
Group work	-0.008 [-0.13, -0.02]	0.052 [-0.01, 0.11]	-0.042 [-0.10, 0.02]	0.056 [0.00, 0.11]	0.059 [0.00, 0.12]	0.027 [-0.03, 0.09]
Age	0.250 [0.20, 0.30]	0.156 [0.11, 0.20]	0.038 [-0.02, 0.09]	0.038 [-0.02, 0.09]	0.051 [-0.01, 0.10]	0.257 [0.21, 0.30]
Gender	0.100 [0.05, 0.15]	-0.005 [-0.06, 0.05]	0.022 [-0.03, 0.08]	-0.048 [-0.10, 0.00]	0.035 [-0.02, 0.09]	0.004 [-0.05, 0.06]
Visual	0.005 [-0.01, 0.11]	0.069 [0.01, 0.13]	0.063 [0.00, 0.12]	-0.024 [-0.08, 0.03]	0.041 [-0.02, 0.10]	0.054 [-0.01, 0.11]
Auditory	0.002 [-0.04, 0.08]	0.073 [0.02, 0.13]	0.023 [-0.03, 0.08]	-0.002 [-0.06, 0.05]	0.013 [-0.04, 0.07]	-0.016 [-0.07, 0.04]
Kinaesthetic	-0.059 [-0.11, 0.00]	-0.124 [-0.18, -0.07]	-0.074 [-0.13, -0.02]	0.022 [-0.03, 0.07]	-0.046 [-0.10, 0.02]	-0.032 [-0.09, 0.02]
	MC self-regulation	Study effort	Study time	Deep learner	Strategic learner	Shallow learner
Study effort	0.594 [0.55, 0.63]					
Study time	0.452 [0.40, 0.49]	0.378 [0.33, 0.43]				
Deep learner	0.431 [0.38, 0.48]	0.360 [0.31, 0.41]	0.285 [0.23, 0.34]			
Strategic learner	-0.213 [-0.26, -0.16]	-0.133 [-0.18, -0.07]	-0.115 [-0.17, -0.06]	-0.791 [-0.81, -0.77]		
Shallow learner	-0.398 [-0.44, -0.35]	-0.394 [-0.44, -0.34]	-0.290 [-0.34, -0.24]	-0.519 [-0.56, -0.48]	-0.103 [-0.15, -0.06]	
Group work	0.113 [0.06, 0.17]	0.094 [0.04, 0.15]	0.084 [0.03, 0.14]	0.020 [-0.03, 0.07]	0.037 [-0.02, 0.09]	-0.081 [-0.14, -0.02]
Age	0.234 [0.18, 0.28]	0.210 [0.16, 0.26]	0.023 [-0.03, 0.08]	0.284 [0.23, 0.33]	-0.200 [-0.25, -0.15]	-0.181 [-0.22, -0.14]
Gender	0.005 [-0.05, 0.06]	0.023 [-0.03, 0.08]	0.086 [0.03, 0.14]	0.086 [0.03, 0.14]	-0.001 [-0.06, 0.06]	-0.130 [-0.18, -0.07]
Visual	0.024 [-0.04, 0.08]	-0.003 [-0.06, 0.06]	0.038 [-0.02, 0.09]	0.067 [0.01, 0.13]	-0.020 [-0.08, 0.04]	-0.089 [-0.15, -0.03]
Auditory	0.065 [0.011, 0.12]	0.039 [-0.01, 0.10]	0.081 [0.03, 0.14]	0.077 [0.02, 0.13]	-0.068 [-0.13, -0.01]	-0.026 [-0.09, 0.03]
Kinaesthetic	-0.078 [-0.13, -0.02]	-0.033 [-0.09, 0.02]	-0.105 [-0.16, -0.05]	-0.126 [-0.18, -0.07]	0.078 [0.02, 0.13]	0.099 [0.04, 0.15]



(2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. <http://dx.doi.org/10.18608/jla.2016.32.20>

	Group work	Age	Gender	Visual	Auditory
Age	-0.022 [-0.08, 0.03]				
Gender	0.026 [-0.03, 0.08]	-0.038 [-0.09, 0.02]			
Visual	0.021 [-0.04, 0.08]	-0.038 [-0.09, 0.02]	-0.046 [-0.10, 0.01]		
Auditory	-0.097 [-0.15, -0.04]	0.025 [-0.03, 0.08]	0.205 [0.15, 0.26]	-0.347 [-0.40, -0.29]	
Kinaesthetic	0.069 [0.01, 0.12]	-0.055 [-0.11, 0.00]	-0.144 [-0.20, -0.09]	-0.541 [-0.58, -0.50]	-0.601 [-0.64, -0.57]

MC: Metacognitive; Intervals are 95% Confidence Intervals based on 1,999 bootstrap samples.

**Table 16: Correlations with GPA by sub-group.**

Sub-group	Mean age	n	Prior academic performance						Personality		Motivation			Self-regulation			Learning Approach			Other			Modality		
			CAO	Maths	English	Method-ical	Human-ities	Applied	Con	Open	SE	EM	IM	SR	StE	StT	Deep	Stra	Shal	Group	Age	Gen	Vis	Aud	Kin
all	23	1207	0.29	0.27	0.17	0.30	0.23	0.17	0.15	0.08	0.12	0.12	0.15	0.13	0.19	0.10	0.23	-0.15	-0.14	-0.08	0.25	0.10	0.05	0.02	-0.05
B-all	21	402	0.35	0.31	0.26	0.30	0.33	0.25	0.21	-0.02	0.08	0.15	0.08	0.11	0.19	0.12	0.18	-0.14	-0.05	-0.03	0.21	-0.02	0.08	-0.04	-0.03
BGen	21	183	0.26	0.23	0.25	0.22	0.26	0.17	0.23	0.09	0.16	0.21	0.12	0.17	0.20	0.10	0.19	-0.08	-0.16	-0.05	0.14	0.10	0.15	0.04	-0.17
BwIT	22	60	0.55	0.52	0.31	0.51	0.34	0.33	0.26	-0.11	-0.01	0.32	-0.03	0.21	0.46	0.13	0.20	-0.16	-0.09	-0.15	0.34	-0.07	-0.05	-0.15	0.16
Blnt	21	64	0.36	0.26	0.09	0.22	0.38	0.15	0.18	-0.38	-0.08	0.09	0.07	0.09	0.12	0.04	0.26	-0.24	-0.07	-0.04	0.29	-0.12	-0.01	-0.05	0.05
Sports	23	95	0.26	0.26	0.30	0.26	0.24	0.22	0.17	0.02	0.05	-0.02	0.05	-0.05	0.05	0.28	0.00	-0.08	0.11	0.05	0.10	0.05	0.04	0.02	-0.06
C-all	24	239	0.18	0.24	-0.03	0.22	0.06	0.10	0.18	0.15	0.12	0.13	0.21	0.14	0.16	0.12	0.29	-0.13	-0.29	-0.09	0.28	0.14	0.09	0.01	-0.08
IT	24	137	-0.02	0.16	-0.27	0.11	-0.12	-0.02	0.28	0.18	0.18	0.18	0.34	0.29	0.21	0.16	0.37	-0.19	-0.32	-0.14	0.42	0.13	0.21	-0.06	-0.12
CDM	23	102	0.09	0.26	-0.08	0.19	-0.01	0.09	0.10	-0.01	0.02	0.18	0.09	0.10	0.21	0.08	0.22	-0.11	-0.19	0.01	0.09	0.00	-0.07	0.08	0.00
E-all	22	172	0.19	0.29	-0.03	0.26	0.14	0.08	0.14	0.10	0.26	0.14	0.29	0.21	0.25	0.11	0.37	-0.28	-0.23	-0.19	0.51	0.04	0.05	0.04	-0.07
EngC	20	73	0.26	0.13	0.03	0.11	0.14	0.20	0.18	0.05	0.21	0.26	0.24	0.22	0.15	0.08	0.36	-0.28	-0.18	0.06	0.41	0.11	0.24	0.10	-0.26
Elec	22	52	0.32	0.44	-0.08	0.43	0.28	0.00	0.35	0.28	0.42	0.21	0.48	0.31	0.35	0.23	0.48	-0.47	-0.24	-0.06	0.39	0.09	0.03	0.01	-0.04
Mech	21	27	0.10	0.30	-0.01	0.21	0.06	-0.04	0.24	0.02	0.42	0.06	0.19	0.15	0.30	0.27	0.20	-0.03	-0.27	-0.16	0.27	0.08	-0.21	0.06	0.14
SECT	27	20	0.23	0.49	0.15	0.38	0.19	0.12	0.08	0.25	0.63	0.34	0.60	0.50	0.55	0.34	0.65	-0.33	-0.58	-0.17	0.68	-0.23	0.06	0.05	-0.11
Hort	28	41	0.07	0.15	-0.09	0.19	0.01	0.24	-0.21	-0.02	0.02	0.09	0.12	0.00	-0.03	-0.15	0.05	-0.15	0.13	-0.14	0.19	0.18	0.05	-0.39	0.30
H-all	25	353	0.25	0.21	0.24	0.26	0.27	0.12	0.05	0.02	0.05	0.09	0.01	0.08	0.15	0.06	0.07	-0.05	-0.04	0.01	0.03	-0.03	0.05	0.00	-0.04
ASC	28	146	0.22	0.28	0.28	0.22	0.23	0.20	-0.09	-0.06	0.04	0.00	-0.08	-0.07	0.07	0.00	0.05	-0.04	-0.03	-0.11	-0.01	0.03	0.02	-0.05	0.02
EC&E	22	80	0.44	0.30	0.37	0.32	0.43	0.23	0.19	0.03	0.09	0.19	0.14	0.18	0.15	0.26	-0.06	0.14	-0.10	-0.04	-0.19	-0.02	-0.07	0.08	-0.03
SCD	25	127	0.27	0.17	0.21	0.31	0.29	0.02	0.14	0.02	0.04	0.14	0.03	0.18	0.22	0.09	0.14	-0.14	-0.03	0.09	0.05	-0.07	0.15	0.01	-0.16
Age [18,23]	20	852	0.44	0.31	0.29	0.38	0.37	0.03	0.14	0.08	0.10	0.12	0.10	0.08	0.14	0.10	0.19	-0.10	-0.13	-0.06	0.12	0.19	0.00	0.05	-0.05
Age [24,28]	25	154	0.16	0.26	0.07	0.26	0.06	0.06	-0.02	-0.04	0.13	0.01	0.04	0.06	0.07	0.10	0.14	-0.14	-0.03	-0.09	0.00	0.08	0.14	-0.02	-0.10
Age [29,60]	38	201	0.17	0.05	0.07	0.00	0.04	0.30	0.06	0.11	0.12	0.13	0.05	0.02	0.19	0.10	0.06	-0.09	0.04	-0.16	-0.08	-0.19	0.12	-0.12	0.01
Male	24	713	0.26	0.27	0.12	0.28	0.20	0.02	0.19	0.14	0.23	0.15	0.25	0.19	0.25	0.12	0.33	-0.23	-0.20	-0.07	0.33		0.09	-0.01	-0.07
Female	23	494	0.30	0.27	0.21	0.29	0.25	0.05	0.10	-0.01	-0.06	0.07	-0.02	0.03	0.08	0.04	0.05	-0.05	0.00	-0.11	0.12		-0.01	0.02	-0.01

Con=Conscientiousness; Open=Openness; SE=Self-efficacy; IM=Intrinsic goal orientation; EM=Extrinsic goal orientation; SR=Metacognitive self-regulation; StE=Study effort; StT=Study time; Deep=Deep learner; Shal=Shallow learner; Stra=Strategic learner; Group=Likes to work in groups; Gen=Gender; Vis=Visual learner; Aud=Auditory learner; Kin=Kinaesthetic learner; B-all=all Business students; BGen=Business General; BwIT=Business with IT; Blnt=International Business; Sport=Sports Management; C=all=all Computing students; IT=Computing(IT); CDM=Creative Digital Media; Eng=all=all Engineering students; EngC=Engineering Common Entry; Elec=Electrical & Computer Engineering; Mech=Mechanics; SECT=Sustainable Electrical and Control Technology; Ho=Horticulture; H-all=all Humanities students; ACS=Applied Social Care; EC&E=Early Childhood & Education; SCD=Social & Community Development.